

Journal of Consulting and Clinical Psychology

JULES D. HOLZBERG, Editor
Wesleyan University

SEYMOUR FISHER, Associate Editor
State University of New York Medical School, Syracuse

ADVISORY EDITORS

SIDNEY J. BLATT
Yale University
LOREN J. CHAPMAN
University of Wisconsin
BERTRAM D. COHEN
Rutgers University
JACOB COHEN
New York University
EMORY L. COWEN
University of Rochester
W. GRANT DAHLSTROM
University of North Carolina
ANTHONY DAVIDS
Brown University
SEYMOUR EPSTEIN
University of Massachusetts
NORMAN L. FARBEROW
Suicide Prevention Center, Los Angeles
AMERIGO FARINA
University of Connecticut
HERMAN FEIFEL
Veterans Administration, Los Angeles
SOL L. GARFIELD
Teachers College, Columbia University
LEONARD D. GOODSTEIN
University of Cincinnati
HARRISON G. GOUGH
University of California, Berkeley
ALFRED B. HEILBRUN
Emory University
IRA ISCOE
University of Texas
DONALD J. KIESLER
Emory University
SHELDON J. KORCHIN
University of California, Berkeley
DAVID LEVINE
University of Nebraska
LEON H. LEVY
Indiana University
MAURICE LORR
Catholic University
BRENDAN A. MAHER
Brandeis University
DAVID MARLOWE
University of California, Santa Cruz

MARTIN MAYMAN
University of Michigan
HERMAN B. MOLISH
Texas Children's Hospital
STUART OSKAMP
Claremont Graduate School
MORRIS B. PARLOFF
National Institute of Mental Health
ALBERT I. RABIN
Michigan State University
RALPH M. REITAN
Indiana University Medical Center
ROBERT ROSENTHAL
Harvard University
ALAN O. ROSS
*State University of New York
at Stony Brook*
JULIAN B. ROTTER
University of Connecticut
EDWIN S. SCHNEIDMAN
National Institute of Mental Health
JEROME L. SINGER
*City College of the City University
of New York*
HANS H. STRUPP
Vanderbilt University
ALEXANDER TOLOR
Fairfield University
READ D. TUDDENHAM
University of California, Berkeley
JOHN L. TURNER
Connecticut Valley Hospital
LEONARD P. ULLMANN
University of Illinois
MORTON WIENER
Clark University
J. RICHARD WITTENBORN
Rutgers University
MELVIN ZAX
University of Rochester
CARL N. ZIMET
*University of Colorado,
School of Medicine*
MARVIN ZUCKERMAN
University of Delaware

HELEN ORR
Managing Editor
HENRY S. SALMON
Advertising and Business Manager

ROBERT D. GORMAN
Subscription Manager
LORRAINE BIROS
Technical Editor

VOLUME 33 1969

PUBLISHED BIMONTHLY BY

• The American Psychological Association, Inc.

Prince and Lemon Sts., Lancaster, Pa. 17604
and 1200 Seventeenth St., N. W., Washington, D. C. 20036

Copyright © 1969 by the American Psychological Association, Inc.

ACKNOWLEDGMENT

The Editor wishes to acknowledge with thanks the advice and consultation provided by the following, who served as special advisory editors on manuscripts considered for publication in Volume 33 of this *Journal*.

Bernard L. Bloom
Joel R. Davitz
Jesse G. Harris
Douglas N. Jackson
Stephen A. Karp
Herbert M. Lefcourt
Joseph M. Masling
Albert Mehrabian
Jum C. Nunnally
Irwin G. Sarason
Paul Satz
Lee B. Sechrest
Ezra Stotland
Bonnie R. Strickland
Roland G. Tharp
William R. Thompson

Contents of Volume 33

<i>Abeles, Norman.</i> See <i>Karl, Norman J.</i>	
<i>Abrahams, Joel.</i> See <i>Spiegel, Don.</i>	
<i>Adams, Jerry, and Canter, Arthur.</i> Performance Characteristics of School Children in the BIP Bender Test	508
<i>Allen, Mildred.</i> See <i>Fagen, Joen.</i>	
<i>Altman, Irwin.</i> See <i>Taylor, Dalmas A.</i>	
<i>Anderson, Lorna M.</i> Personality Characteristics of Parents of Neurotic, Aggressive, and Normal Preadolescent Boys	575✓
<i>Arbuthnot, Jack, and Gruenfeld, Leopold.</i> Field Independence and Educational-Vocational Interests	631
<i>Ashem, Beatrice A.</i> Use of Speech in the Inhibitory Control of Motoric Behavior of Schizophrenic Patients	226
<i>Barbee, Margaret S., Berry, K. L., and Micek, Leo A.</i> Relationship of Work Therapy to Psychiatric Length of Stay and Readmission	735
<i>Barber, Theodore Xenophon.</i> Invalid Arguments, Postmortem Analyses, and the Experimenter Bias Effect	11
<i>Barber, Theodore Xenophon, Calverley, David S., Forgione, Albert, McPeake, John D., Chaves, John F., and Bowen, Barbara.</i> Five Attempts to Replicate the Experimenter Bias Effect	1
<i>Barclay, Allan, and Yater, Allan C.</i> Comparative Study of the Wechsler Preschool and Primary Scale of Intelligence and the Stanford-Binet Intelligence Scale, Form L-M, among Culturally Deprived Children	257✓
<i>Berry, K. L.</i> See <i>Barbee, Margaret S.</i>	
<i>Berry, K. L., and Miskimins, R. W.</i> Concept of Self and Posthospital Vocational Adjustment	103
<i>Berzins, Juris, I.</i> See <i>Dublin, James E.</i>	
<i>Berzins, Juris I., and Seidman, Edward.</i> Differential Therapeutic Responding of A and B Quasi-Therapists to Schizoid and Neurotic Communications	279
<i>Blackburn, R.</i> Sensation Seeking, Impulsivity, and Psychopathic Personality	571✓
<i>Blatt, Sidney J.</i> See <i>Fox, Elizabeth.</i>	
<i>Bolton, Laura Specht.</i> See <i>Cattell, Raymond B.</i>	
<i>Bowen, Barbara.</i> See <i>Barber, Theodore Xenophon.</i>	
<i>Brown, Oliver H.</i> See <i>Veldman, Donald J.</i>	
<i>Broughton, Elizabeth.</i> See <i>Fagin, Joen.</i>	
<i>Bugle, Charles.</i> See <i>Rutner, Ivan Toby</i>	
<i>Burn, Julian.</i> See <i>Parsons, Oscar A.</i>	
<i>Bush, Marshall, Hatcher, Robert, and Mayman, Martin.</i> Reality Attentiveness-Inattentiveness and Externalization-Internalization in Defensive Style	343
<i>Buttiglieri, Matthew W., Woodson, M. I. Charles E., Geuente, Marie, and Thomson, Mae.</i> Driver Accidents and the Neuropsychiatric Patient	381✓
<i>Calverley, David S.</i> See <i>Barber, Theodore Xenophon.</i>	
<i>Callens, Charlotte, J., and Meltzer, Malcolm L.</i> Effect of Intelligence, Anxiety, and Diagnosis on Arithmetic and Digit Span Performance on the WAIS	630
<i>Cancro, Robert.</i> Abstraction on Proverbs in Process-Reactive Schizophrenia	267
<i>Cannon, John R., and Carkhuff, Robert R.</i> Effects of Rater Level of Functioning and Experience upon the Discrimination of Facilitative Conditions	189
<i>Canter, Arthur.</i> See <i>Adams, Jerry.</i>	
<i>Carkhuff, Robert R.</i> See <i>Cannon, John R.</i>	
<i>Carlson, Rae.</i> Rorschach Prediction of Success in Clinical Training: A Second Look	699

<i>Cattell, Raymond B., and Bolton, Laura Specht.</i> What Pathological Dimensions Lie beyond the Normal Dimensions of the 16 PF? A Comparison of MMPI and 16 PF Factor Domains	18
<i>Chapman, Loren J.</i> Schizomimetic Conditions and Schizophrenia	646
<i>Chaves, John F.</i> See <i>Barber, Theodore Xenophon.</i>	
<i>Clark, Betty.</i> See <i>Fagan, Joen.</i>	
<i>Cohen, David Benjamin.</i> Frequency of Dream Recall Estimated by Three Methods and Related to Defense Preference and Anxiety	661
<i>Conn, Lane K.</i> See <i>Tulkin, Steven R.</i>	
<i>Corey, Michel.</i> See <i>Klein, Edward B.</i>	
<i>Cottle, Thomas J.</i> Temporal Correlates of the Achievement Value and Manifest Anxiety	541
<i>Cowden, James E.</i> Prediction Enhancement through the Use of Moderator Variables	621
<i>Cowden, James E., and Pacht, Asher R.</i> Relationship of Selected Psychosocial Variables to Prognostic Judgments	254
<i>Cromwell, Rue L.</i> See <i>Neale, John M.</i>	
<i>Cross, Herbert J.</i> College Students' Memories of Their Parents: A Factor Analysis of the CRPBI	275
<i>Cudrin, Jay M.</i> Intelligence of Volunteers as Research Subjects	501
<i>Danet, Burton N.</i> Impact of Audio-Visual Feedback on Group Psychotherapy	632
<i>Darbonne, Allen R.</i> Study of Psychological Content in the Communications of Suicidal Individuals	590
<i>Darbonne, Allen R.</i> Suicide and Age: A Suicide Note Analysis	46
<i>David, Anthony.</i> Ego Functions in Disturbed and Normal Children: Aspiration, Inhibition, Time Estimation, and Delayed Gratification	61
<i>Davison, Gerald C.</i> See <i>Wilson, G. Terence</i>	
<i>Detre, Thomas.</i> See <i>Harrow, Martin.</i>	
<i>Dicken, Charles.</i> Predicting the Success of Peace Corps Community Development Workers	597
<i>Dickstein, Louis S.</i> Prospective Span as a Cognitive Ability	757
<i>Dittmann, Allen T.</i> See <i>Klein, Marjorie H.</i>	
<i>Domino, George.</i> Maternal Personality Correlates of Son's Creativity	180
<i>Doster, Joseph A., and Strickland, Bonnie Ruth.</i> Perceived Child-Rearing Practices and Self-Disclosure Patterns	382
<i>Dublin, James E., Elton, Charles F., and Berzins, Juris I.</i> Some Personality and Aptitudinal Correlates of the "A-B" Therapist Scale	739
<i>Ebner, Eugene, and Ritzler, Barry.</i> Perceptual Recognition in Chronic and Acute Schizophrenics	200
<i>Efran, Jay S., and Korn, Paul R.</i> Measurement of Social Caution: Self-Appraisal, Role Playing, and Discussion Behavior	78
<i>Eisenman, Russell.</i> Creativity, Awareness, and Liking	157
<i>Eisenthal, Sherman, and Sherman, Lewis J.</i> Psychological Characteristics of Neighborhood Youth Corps Enrollees	420
<i>Elton, Charles F.</i> See <i>Dublin, James E.</i>	
<i>Emerson, Patricia.</i> See <i>Fagan, Joen.</i>	
<i>Evans, Ray B.</i> Childhood Parental Relationships of Homosexual Men	129
<i>Fagan, Joen, Broughton, Elizabeth, Allen, Mildred, Clark, Betty, and Emerson, Patricia.</i> Comparison of the Binet and WPPSI with Lower-Class Five-Year-Olds	607
<i>Fancher, Raymond E.</i> Group and Individual Accuracy in Person Perception	127
<i>Farina, Amerigo.</i> See <i>Ring, Kenneth</i>	

<i>Fazio, Anthony F.</i> Verbal and Overt-Behavioral Assessment of a Specific Fear . . .	705
<i>Feifel, Herman.</i> Attitudes toward Death: A Psychological Perspective . . .	292
<i>Ferrante, Anthony.</i> See <i>Harrow, Martin.</i>	
<i>Fischer, Constance T.</i> Intelligence Defined as Effectiveness of Approaches . . .	668
<i>Fischer, Edward H., and Winer, David.</i> Participation in Psychological Research: Relation to Birth Order and Demographic Factors . . .	610
<i>Fontana, Alan F., and Gessner, Theodore.</i> Patients' Goals and the Manifestation of Psychopathology . . .	247
<i>Forgione, Albert.</i> See <i>Barber, Theodore Xenophon.</i>	
<i>Foss, Adeline.</i> See <i>Ravensborg, Milton R.</i>	
<i>Fox, David A.</i> See <i>Harrow, Martin.</i>	
<i>Fox, Elizabeth, and Blatt, Sidney J.</i> An Attempt to Test Assumptions about Some Indications of Negativism on Psychological Tests . . .	365
<i>Friar, John T.</i> See <i>Mehrabian, Albert.</i>	
<i>Friend, Ronald.</i> See <i>Watson, David.</i>	
<i>Fuller, George D., and Kates, Solis L.</i> Word Association Repertoires of Schizophrenics and Normals . . .	497
<i>Geisinger, David.</i> See <i>Murray, Edward J.</i>	
<i>Gessner, Theodore.</i> See <i>Fontana, Alan F.</i>	
<i>Giebink, John W., and Stover, Donald O.</i> Adjustment, Mental Health Opinions, and Proficiency of Child Care Personnel . . .	532
<i>Gill, Merton M.</i> See <i>Klein, Marjorie H.</i>	
<i>Gleser, Goldine C., and Ihlevich, David.</i> An Objective Instrument for Measuring Defense Mechanisms . . .	51
<i>Golland, Jeffrey H.</i> See <i>Herrell, James M.</i>	
<i>Gorlow, Leon.</i> See <i>Knestrick, James L.</i>	
<i>Gould, Laurence J.</i> See <i>Klein, Edward B.</i>	
<i>Granlund, Elnore, and Knowles, Lois.</i> Child-Parent Identification and Academic Underachievement . . .	495
<i>Greenberg, Irene.</i> See <i>Mosher, Donald L.</i>	
<i>Greenberg, Roger P.</i> Effects of Presession Information on Perception of the Therapist and Receptivity to Influence in a Psychotherapy Analogue . . .	425
<i>Gruenfeld, Leopold.</i> See <i>Arbuthnot, Jack.</i>	
<i>Guenette, Marie.</i> See <i>Buttiglieri, Matthew W.</i>	
<i>Gundlach, Ralph H.</i> Childhood Parental Relationships and the Establishment of Gender Roles of Homosexuals . . .	136
<i>Hagebak, Robert W., and Parker, George V. C.</i> Therapist Directiveness, Client Dominance, and Therapy Resistance . . .	536
<i>Hagey, Joan.</i> See <i>Ziller, Robert C.</i>	
<i>Hall, Janet.</i> See <i>MacDonald, A. P., Jr.</i>	
<i>Hall, Leon P., and LaDriere, LaVerne.</i> Patterns of Performance on WISC Similari- ties in Emotionally Disturbed and Brain-Damaged Children . . .	357
<i>Halverson, Charles F., Jr., and Shore, Roy E.</i> Self-Disclosure and Interpersonal Functioning . . .	213
<i>Hammer, Emanuel F.</i> DAP: Back against the Wall . . .	151
<i>Hanson, Barbara A.</i> See <i>Magaro, Peter A.</i>	
<i>Harris, Sandra.</i> See <i>Masling, Joseph.</i>	
<i>Harrow, Martin, and Ferrante, Anthony.</i> Locus of Control in Psychiatric Patients .	582
<i>Harrow, Martin, Fox, David A., and Detre, Thomas.</i> Self-Concept of the Married Psychiatric Patient and His Mate's Perception of Him . . .	235
<i>Hatcher, Robert.</i> See <i>Bush, Marshall.</i>	

<i>Hathaway, Starke R., Reynolds, Phyllis C., and Monachesi, Elio D.</i> Follow-Up of the Later Careers and Lives of 1,000 Boys who Dropped Out of High School . . .	370
<i>Hathaway, Starke R., Reynolds, Phyllis C., and Monachesi, Elio D.</i> Follow-Up of 812 Girls 10 Years after High School Dropout . . .	383
<i>Heller, Kenneth.</i> Effects of Modeling Procedures in Helping Relationships . . .	522
<i>Herrell, James M., and Golland, Jeffrey, H.</i> Should WISC Subjects Explain Picture Arrangement Stories? . . .	761
<i>Hersch, Paul D., Kulik, James A., and Scheibe, Karl E.</i> Personal Characteristics of College Volunteers in Mental Hospitals . . .	30
<i>Hocking, John, and Robertson, Malcolm.</i> Sensation Seeking Scale as a Predictor of Need for Stimulation during Sensory Restriction . . .	367
<i>Hodges, William F., and Spielberger, Charles D.</i> Digit Span: An Indicant of Trait or State Anxiety? . . .	430
<i>Hoffnung, Robert J.</i> Conditioning and Transfer of Affective Self-References in a Role-Played Counseling Interview . . .	527
<i>Hogan, Robert.</i> Development of an Empathy Scale . . .	307
<i>Holmes, Douglas S.</i> Sensing Humor: Latency and Amplitude of Response Related to MMPI Profiles . . .	296
<i>Hooker, Evelyn.</i> Parental Relations and Male Homosexuality in Patient and Non-patient Samples . . .	140
<i>Houts, Peter S., MacIntosh, Shirley, and Moos, Rudolf H.</i> Patient-Therapist Interdependence: Cognitive and Behavioral . . .	40
<i>Howard, Kenneth I., Krause, Merton S., and Orlinsky, David E.</i> Direction of Affective Influence in Psychotherapy . . .	614
<i>Hsieh, Theodore Tin-Yee, Shybut, John, and Lotsof, Erwin J.</i> Internal versus External Control and Ethnic Group Membership: A Cross-Cultural Comparison . . .	122
<i>Hurley, John R.</i> See <i>Melnick, Barry.</i>	
<i>Ihilevich, David.</i> See <i>Gleser, Goldine C.</i>	
<i>Jacobs, Martin A.</i> See <i>Spilken, Aron Z.</i>	
<i>Kalter, Neil.</i> See <i>Marsden, Gerald.</i>	
<i>Karl, Norman J., and Abeles, Norman.</i> Psychotherapy Process as a Function of the Time Segment Sampled . . .	207
<i>Kates, Solis L.</i> See <i>Fuller, George D.</i>	
<i>Keith-Spiegel, Patricia.</i> See <i>Spiegel, Don.</i>	
<i>Kirtley, Donald D., and Sacks, Joseph M.</i> Reactions of a Psychotherapy Group to Ambiguous Circumstances Surrounding the Death of A Group Member . . .	195
<i>Klein, Edward B., Gould, Laurence J., and Corey, Michel.</i> Social Desirability in Children: An Extension and Replication . . .	128
<i>Klein, Marjorie H., Dittmann, Allen T., Parloff, Morris B., and Gill, Merton M.</i> Behavior Therapy: Observations and Reflections . . .	259
<i>Knestrick, James L., and Gorlow, Leon.</i> Effect of Information, Subject Age, and Instructional Set on Predictive Judgments . . .	691
<i>Knitzer, Jane.</i> See <i>Spilken, Aron Z.</i>	
<i>Knowles, Lois.</i> See <i>Granlund, Elnore.</i>	
<i>Korn, Paul R.</i> See <i>Efran, Jay S.</i>	
<i>Kranitz, Lionel.</i> See <i>Spiegel, Don.</i>	
<i>Krapf, Jon E., and Nawas, M. Mike.</i> Client-Therapist Relationship Factor in Systematic Desensitization . . .	435
<i>Krause, Merton S.</i> See <i>Howard, Kenneth I.</i>	
<i>Kulik, James A.</i> See <i>Hersch, Paul D.</i>	

<i>Kurtz, Richard M.</i> Sex Differences and Variations in Body Attitudes	625
<i>Kushner, Estelle N.</i> See <i>Taylor, Dalmis A.</i>	
<i>LaDriere, LaVerne.</i> See <i>Hall, Leon P.</i>	
<i>Lamb, Nancy.</i> See <i>Leff, Stephen.</i>	
<i>Lefcourt, Herbert M.</i> Need for Approval and Threatened Negative Evaluation as Determinants of Expressiveness in a Projective Test	96
<i>Leff, Stephen, and Lamb, Nancy.</i> Experimental Approach to Defining the Role of Social Desirability in Personality Assessment: Is There One Response Process or Two?	287
<i>Levitz, Leonard S., and Ullmann, Leonard P.</i> Manipulation of Indications of Disturbed Thinking in Normal Subjects	633
<i>Levy, Leon H.</i> Reflections on Replications and the Experimenter Bias Effect	15
<i>Liebert, Robert M.</i> See <i>Morris, Larry W.</i>	
<i>Logue, Patrick E.</i> See <i>Watson, Charles G.</i>	
<i>Long, Barbara H.</i> See <i>Ziller, Robert C.</i>	
<i>Lotsof, Erwin J.</i> See <i>Hsieh, Theodore Tin-Yee.</i>	
<i>MacDonald, A. P., Jr., and Hall, Janet.</i> Perception of Disability by the Nondisabled	654
<i>MacIntosh, Shirley.</i> See <i>Houts, Peter S.</i>	
<i>Magaro, Peter A., and Hanson, Barbara A.</i> Perceived Maternal Nurture and Control of Process Schizophrenics, Reactive Schizophrenics, and Normals	507
<i>Marsden, Gerald, and Kalter, Neil.</i> Bodily Concerns and the WISC Object Assembly Subtest	391
<i>Masling, Joseph, and Harris, Sandra.</i> Sexual Aspects of TAT Administration	166
<i>Mayman, Martin.</i> See <i>Bush, Marshall.</i>	
<i>McPeake, John D.</i> See <i>Barber, Theodore Xenophon.</i>	
<i>Mehrabian, Albert, and Friar John T.</i> Encoding of Attitude by A Seated Communicator via Posture and Position Cues	330
<i>Meichenbaum, Donald H.</i> Validation Study of Marital Status and the Self-Report Scale for Process-Reactive Schizophrenia	351
<i>Melnick, Barry, and Hurley, John R.</i> Distinctive Personality Attributes of Child-Abusing Mothers	746
<i>Meltzer, Malcolm L.</i> See <i>Callens, Charlotte J.</i>	
<i>Mendelsohn, Gerald A.</i> See <i>Pittel, Stephen M.</i>	
<i>Micek, Leo A.</i> See <i>Barbee, Margaret S.</i>	
<i>Miskimins, R. W.</i> See <i>Berry, K. L.</i>	
<i>Miskimins, R. W., and Wilson, Lowell T.</i> Revised Suicide Potential Scale	258
<i>Monachesi, Elio D.</i> See <i>Hathaway, Starke R.</i>	
<i>Mondani, Michele S., and Tutko, Thomas A.</i> Relationship of Academic Underachievement to Incidental Learning	558
<i>Moos, Rudolf H.</i> See <i>Houts, Peter S.</i>	
<i>Morris, Larry W., and Liebert, Robert M.</i> Effects of Anxiety on Timed and Untimed Intelligence Tests: Another Look	240
<i>Mosher, Donald L., and Greenberg, Irene.</i> Females' Affective Responses to Reading Erotic Literature	472
<i>Muller, James J.</i> See <i>Spilken, Aron Z.</i>	
<i>Muller, John P.</i> See <i>Tulkin, Steven R.</i>	
<i>Murray, Edward J., Seagull, Arthur, and Geisinger, David.</i> Motivational Patterns in the Families of Adjusted and Maladjusted Boys	337
<i>Nawas, M. Mike.</i> See <i>Krapfl, Jon E.</i>	
<i>Neale, John M., and Cromwell, Rue L.</i> Preference for Complexity in Acute Schizophrenics	245

<i>Neuringer, Charles.</i> See <i>Van De Mark, Steven N.</i>	
<i>Nickels, James B.</i> See <i>Williams, Christene Blanton.</i>	
<i>Norman, Russell P.</i> Extreme Response Tendency as a Function of Emotional Adjustment and Stimulus Ambiguity	406
<i>Orlinsky, David E.</i> See <i>Howard, Kenneth I.</i>	
<i>Owen, David R.</i> See <i>Sines, Jacob O.</i>	
<i>Pacht, Asher R.</i> See <i>Cowden, James E.</i>	
<i>Parker, George V. C.</i> See <i>Hagebak, Robert W.</i>	
<i>Parloff, Morris B.</i> See <i>Klein, Marjorie H.</i>	
<i>Parsons, Oscar A., Vega, Arthur, Jr., and Burn, Julian.</i> Different Psychological Effects of Lateralized Brain Damage.	551
<i>Patterson, John R.</i> See <i>Wilson, Glenn D.</i>	
<i>Pauker, Jerome D.</i> See <i>Sines, Jacob O.</i>	
<i>Pearson, Douglas.</i> Object-Discrimination Learning Set Acquisition in Young Cerebral Palsied Children in Relation to Tested and Rated Intelligence . . .	478
<i>Pettit, Tupper F.</i> Analilty and Time	170
<i>Pittel, Stephen M., and Mendelsohn, Gerald A.</i> Situational Appraisal Inventory: Development and Validation of a Measure of Evaluative Attitudes	396
<i>Rardin, Max W.</i> Treatment of a Phobia by Partial Self-Desensitization: A Case Study	125
<i>Ravensborg, Milton R., and Foss, Adeline.</i> Suicide and Natural Death in a State Hospital Population: A Comparison of Admission Complaints, MMPI Profiles, and Social Competence Factors	466
<i>Reynolds, Phyllis C.</i> See <i>Hathaway, Starke R.</i>	
<i>Ring, Kenneth, and Farina, Amerigo.</i> Personal Adjustment as a Determinant of Aggressive Behavior toward the Mentally Ill	683
<i>Ritzler, Barry.</i> See <i>Ebner, Eugene.</i>	
<i>Roberts, Lynn K.</i> See <i>Sullivan, Patrick F.</i>	
<i>Robertson, Malcolm.</i> See <i>Hocking, John.</i>	
<i>Rosenbaum, Gerald.</i> Schizophrenia as a "Put-On"	642
<i>Rosenthal, Robert.</i> On Not So Replicated Experiments and Not So Null Results . .	7
<i>Rutner, Ivan Toby, and Bugle, Charles.</i> An Experimental Procedure for the Modification of Psychotic Behavior	651
<i>Sabatino, David A.</i> Identifying Neurologically Impaired Children through a Test of Auditory Perception	184
<i>Sacks, Joseph M.</i> See <i>Kirtley, Donald D.</i>	
<i>Sattler, Jerome M.</i> Effects of Cues and Examiner Influence on Two Wechsler Subtests	716
<i>Scheibe, Karl E.</i> See <i>Hersch, Paul D.</i>	
<i>Schwitzgebel, Robert L.</i> Preliminary Socialization for Psychotherapy of Behavior-Disordered Adolescents	71
<i>Seagull, Arthur.</i> See <i>Murray, Edward J.</i>	
<i>Seidman, Edward.</i> See <i>Berzins, Juris I.</i>	
<i>Sherman, Lewis J.</i> See <i>Eisenthal, Sherman.</i>	
<i>Shore, Roy E.</i> See <i>Halverson, Charles F., Jr.</i>	
<i>Shybut, John.</i> See <i>Hsieh, Theodore Tin-Yee.</i>	
<i>Sines, Jacob O., Pauker, Jerome D., Sines, Lloyd K., and Owen, David R.</i> Identification of Clinically Relevant Dimensions of Children's Behavior	728
<i>Sines, Lloyd K.</i> See <i>Sines, Jacob O.</i>	
<i>Smith, Gene M.</i> Relations between Personality and Smoking Behavior in Preadult Subjects	710

<i>Smith, Jean J.</i> Psychiatric Hospital Experience and Attitudes toward "Mental Illness"	302
<i>Smith, Mary Dell C.</i> See <i>Ziller, Robert C.</i>	
<i>Smith, Ronald E.</i> Motivational Properties of Personality Conflict	506
<i>Spiegel, Don.</i> SPI Discriminators among Four Psychological Health-Sickness Levels	750
<i>Spiegel, Don, Keith-Spiegel, Patricia, Abrahams, Joel, and Kranitz, Lionel.</i> Humor and Suicide: Favorite Jokes of Suicidal Patients	504
<i>Spielberger, Charles D.</i> See <i>Hodges, William F.</i>	
<i>Spilken, Aron Z., Jacobs, Martin A., Muller, James J., and Knitzer, Jane.</i> Personality Characteristics of Therapists: Description of Relevant Variables and Examination of Conscious Preferences	317
<i>Stover, Donald O.</i> See <i>Giebkink, John W.</i>	
<i>Streissguth, Ann Pytkowicz, Wagner, Nathaniel N., and Weschler, John C.</i> Effects of Sex, Illness, and Hospitalization on Daydreaming	218
<i>Strickland, Bonnie Ruth.</i> See <i>Doster, Joseph A.</i>	
<i>Stuart, Richard B.</i> Operant-Interpersonal Treatment for Marital Discord	675
<i>Sullivan, Patrick, F., and Roberts, Lynn K.</i> Relationship of Manifest Anxiety to Repression-Sensitization on the MMPI	763
<i>Taylor, Dalmás, A. Altman, Irwin, Wheeler, Ladd, and Kushner, Estelle N.</i> Personality Factors Related to Response to Social Isolation and Confinement	411
<i>Thelen, Mark H.</i> Repression-Sensitization: Its Relation to Adjustment and Seeking Psychotherapy among College Students	161
<i>Thomson, Mae.</i> See <i>Buttiglieri, Matthew W.</i>	
<i>Towbin, Alan P.</i> Self-Care Unit: Some Lessons in Institutional Power	561
<i>Truax, Charles B., and Wargo, Donald G.</i> Effects of Vicarious Therapy Pretraining and Alternate Sessions on Outcome in Group Psychotherapy with Outpatients	440
<i>Tulkin, Steven R., Muller, John P., and Conn, Lane K.</i> Need for Approval and Popularity: Sex Differences in Elementary School Students	35
<i>Tutko, Thomas A.</i> See <i>Mondani, Michele S.</i>	
<i>Ullmann, Leonard P.</i> See <i>Levitz, Leonard S.</i>	
<i>Van De Mark, Steven N., and Neuringer, Charles.</i> Effect of Physical and Cognitive Somatic Arousal on Rorschach Responses: An Experimental Test of the Assumption that Body Image Influences the Perceptual Organization of Unstructured Stimuli	458
<i>Vega, Arthur, Jr.</i> See <i>Parsons, Oscar A.</i>	
<i>Veldman, Donald J., and Bown, Oliver H.</i> Personality and Performance Characteristics Associated with Cigarette Smoking among College Freshmen	109
<i>Vesprani, George J.</i> Personality Correlates of Accurate Empathy in a College Companion Program	722
<i>Vestre, Norris D., and Zimmermann, Robert.</i> Validity of Informants' Ratings of the Behavior and Symptoms of Psychiatric Patients	175
<i>Wagner, Nathaniel N.</i> See <i>Streissguth, Ann Pytkowicz.</i>	
<i>Wanderer, Zev William.</i> Validity of Clinical Judgments Based on Human Figure Drawings	143
<i>Wargo, Donald G.</i> See <i>Truax, Charles B.</i>	
<i>Watson, Charles G., and Logue, Patrick E.</i> Interrelationships of Several Process-Reactive Measures	120
<i>Watson, David, and Friend, Ronald.</i> Measurement of Social-Evaluative Anxiety	448
<i>Weschler, John C.</i> See <i>Streissguth, Ann Pytkowicz.</i>	
<i>Whalen, Carol.</i> Effects of a Model and Instructions on Group Verbal Behaviors	509
<i>Wheeler, Ladd.</i> See <i>Taylor, Dalmás A.</i>	

Williams, Christene Blanton, and Nickels, James B. Internal-External Control Dimension as Related to Accident and Suicide Proneness 485

Wilson, G. Terence, and Davison, Gerald C. Aversion Techniques in Behavior Therapy: Some Theoretical and Metatheoretical Considerations 327

Wilson, Glenn D., and Patterson, John R. Conservatism as a Predictor of Humor Preferences 271

Wilson, Lowell T. See *Miskimins, R. W.*

Winer, David. See *Fischer, Edward H.*

Woodson, M. I. Charles E. See *Buttiglieri, Matthew W.*

Yater, Allan C. See *Barclay, Allan*

Ziller, Robert C., Hagey, Joan, Smith, Mary Dell C., and Long, Barbara H. Self-Esteem: A Self-Social Construct 84

Zimmermann, Robert. See *Vestre, Norris D.*

CONSULTING PSYCHOLOGIST

Mature Ph.D. Psychologist with strong clinical skills. Should be experienced counseling with mature adults. Supervisory experience desirable. Prefer knowledge of guidance and industrial counseling.

To counsel executives and management personnel in areas of personal development and career planning. Handle assignments in assessment, training and organization.

Starting salary \$15,000 to \$25,000 based upon experience and qualifications. Attractive benefits, growth opportunity. Little or no travel.

Opportunities in New York, Chicago, Dallas, Los Angeles, San Francisco and other major cities.

For consideration phone or submit resume in confidence:

C. T. Deeble, Ph.D.

American Management Psychologists

**200 South Michigan Avenue, Chicago, Illinois 60604
312-427-9335**

Journal of **APPLIED BEHAVIOR ANALYSIS**

MONTROSE M. WOLF, *Editor*
T. AYLLON, *Associate Editor*
DONALD M. BAER, *Associate Editor*
HAROLD LEITENBERG, *Associate Editor*
O. IVAR LOVAAS, *Associate Editor*
VICTOR G. LATIES, *Executive Editor*

EDITORIAL BOARD

W. S. AGRAS			M. SIDMAN
N. H. AZRIN	I. GOLDDIAMOND	O. R. LINDSLEY	G. M. SIEGEL
A. BANDURA	J. G. HOLLAND	J. L. MICHAEL	B. F. SKINNER
W. C. BECKER	B. L. HOPKINS	G. R. PATTERSON	H. N. SLOANE
J. S. BIRNBAUER	F. S. KELLER	T. R. RISLEY	J. E. SPRADLIN
C. B. FERSTER	P. J. LANG	J. A. SHERMAN	A. W. STAATS

Fall, 1969 (Vol. 2, No. 3)

O. Ivar Lovaas and James Q. Simmons, "Manipulation of self-destruction in three retarded children."

Robert G. Wahler, "Oppositional children: a quest for parental reinforcement control."

Gilbert W. Schmidt and Roger E. Ulrich, "Effects of group-contingent events upon classroom noise."

Barbara H. Wasik, Kathryn Senn, Roberta H. Welch and Barbara R. Cooper, "Behavior modification with culturally-deprived school children: two case studies."

K. A. Lattal, "Contingency management of toothbrushing behavior in a summer camp for children."

Arthur D. Colman and John J. Boren, "An information system for measuring patient behavior and its use by staff."

Leija V. McReynolds, "Application of time-out from positive reinforcement for increasing the efficiency of speech training."

Robert P. Cantrell, Mary Lynn Cantrell, Clifton M. Huddleston and Ralph L. Wooldridge, "Contingency contracting with school problems." (discussion article)

R. M. Sanders, B. L. Hopkins, and M. B. Walker, "An inexpensive method for making data records of complex behavior." (technical note)

Address subscription orders to Mary L. Sherman,
Department of Human Development, University of Kansas,
Lawrence, Kansas 66044

Students \$4 • Individuals (personal use) \$8 • Institutions \$16

JOURNAL OF CONSULTING and CLINICAL PSYCHOLOGY

February 1969

Vol. 33, No. 1

Contents

ARTICLES

Five Attempts to Replicate the Experimenter Bias Effect: <i>Theodore Xenophon Barber, David S. Calverley, Albert Forgione, John D. McPeake, John F. Chaves, and Barbara Bowen</i>	1
On Not So Replicated Experiments and Not So Null Results: <i>Robert Rosenthal</i> ...	7
Invalid Arguments, Postmortem Analyses, and the Experimenter Bias Effect: <i>Theodore Xenophon Barber</i>	11
Reflections on Replications and the Experimenter Bias Effect: <i>Leon H. Levy</i>	15
What Pathological Dimensions Lie beyond the Normal Dimensions of the 16 PF? A Comparison of MMPI and 16 PF Factor Domains: <i>Raymond B. Cattell and Laura Specht Bolton</i>	18
Personal Characteristics of College Volunteers in Mental Hospitals: <i>Paul D. Hersch, James A. Kulik, and Karl E. Scheibe</i>	30
Need for Approval and Popularity: Sex Differences in Elementary School Students: <i>Steven R. Tulkin, John P. Muller, and Lane K. Conn</i>	35
Patient-Therapist Interdependence: Cognitive and Behavioral: <i>Peter S. Houts, Shirley MacIntosh, and Rudolf H. Moos</i>	40
Suicide and Age: A Suicide Note Analysis: <i>Allen R. Darbonne</i>	46
An Objective Instrument for Measuring Defense Mechanisms: <i>Goldine C. Gleser and David Ihilevich</i>	51
Ego Functions in Disturbed and Normal Children: Aspiration, Inhibition, Time Estimation, and Delayed Gratification: <i>Anthony David</i>	61
Preliminary Socialization for Psychotherapy of Behavior-Disordered Adolescents: <i>Robert L. Schwitzgebel</i>	71
Measurement of Social Caution: Self-Appraisal, Role Playing, and Discussion Behavior: <i>Jay S. Efran and Paul R. Korn</i>	78
Self-Esteem: A Self-Social Construct: <i>Robert C. Ziller, Joan Hagey, Mary Dell C. Smith, and Barbara H. Long</i>	84
Need for Approval and Threatened Negative Evaluation as Determinants of Expressiveness in a Projective Test: <i>Herbert M. Lefcourt</i>	96
Concept of Self and Posthospital Vocational Adjustment: <i>K. L. Berry and R. W. Miskimins</i>	103
Personality and Performance Characteristics Associated with Cigarette Smoking among College Freshmen: <i>Donald J. Veldman and Oliver H. Bown</i>	109

NOTES AND COMMENTS

Interrelationships of Several Process-Reactive Measures: <i>Charles G. Watson and Patrick E. Logue</i>	120
Internal versus External Control and Ethnic Group Membership: A Cross-Cultural Comparison: <i>Theodore Tin-Yee Hsieh, John Shybut, and Erwin J. Lofof</i>	122
Treatment of a Phobia by Partial Self-Desensitization: A Case Study: <i>Max W. Rardin</i>	125

BRIEF REPORTS

Group and Individual Accuracy in Person Perception: <i>Raymond E. Fancher</i>	127
Social Desirability in Children: An Extension and Replication: <i>Edward B. Klein, Laurence J. Gould, and Michel Corey</i>	128

LIST OF MANUSCRIPTS.....	70, 77, 95
--------------------------	------------

The *Journal of Consulting and Clinical Psychology* will accept Brief Reports of research studies in clinical psychology for early publication without expense to the author. The procedure is intended to permit the publication of soundly designed studies of specialized interest or limited importance which cannot now be accepted because of lack of space. Several pages in each issue will be devoted to Brief Reports, published in the order of their receipt without respect to the dates of receipt of the regular articles. Most Brief Reports appear in the first or second issue to go to press following their final acceptance.

An author who wishes to submit a Brief Report:

1. Sends the Brief Report, limited to one printed page and prepared according to the specifications given below.
2. Also sends to the Editor a full report of the research study, in sufficient detail to give a clear account of its background, procedure, results, and conclusions, which will be filed with the National Auxiliary Publications Service (NAPS) to insure indefinite availability.
3. Prepares at least 100 mimeographed copies of the full report, which the author will send without charge to all who request it as long as the supply lasts.
4. Agrees not to submit the full report to another journal of general circulation.

Specifications

Brief Report. The Brief Report should give a clear, condensed summary of the procedure of the study and as full an account of the results as space permits.

To insure that the Brief Report will be no longer than one printed page, its typescript, in-

cluding all matter except the title and the author's lines, must not exceed 85 lines averaging 42 characters and spaces in length. Set the type-writer margins for short lines of 42 characters, which are 3.5 inches long in elite typing, and 4.2 inches long in pica.

The manuscript of the Brief Report must be *double spaced* throughout. Except for its short lines, it follows the standard style of the 1967 revision of the *APA Publication Manual*. Headings, tables, and references are avoided or, if essential, must be counted in the 85 lines. Each Brief Report must be accompanied by a footnote in the style below, which is typed on a separate sheet and *not* counted in the 85-line quota:

¹ An extended report of this study may be obtained without charge from John Doe (giving the author's full name and address) or for a fee from the National Auxiliary Publications Service. Order NAPS Document — from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001; remitting \$ — for microfiche or \$ — for photocopies.

Extended report. Because the extended report is intended for photoduplication, and is not copy to be sent to a printer, its style should differ in several ways from that of other manuscripts: (a) The extended report should be typed with single spacing for economy in duplication. (b) Tables and figures should be placed adjacent to the text which refers to them. A caption should be typed below each figure. (c) Footnotes should be typed at the bottom of the page on which reference is made to them. In other respects, the full report is prepared in the style specified by the *Publication Manual*.

Style of Manuscripts. Manuscripts must be prepared in the style described in the *Publication Manual of the American Psychological Association (1967 Revision)*, obtainable for \$1.50 from the Washington office of the APA.

Early Publication. Authors may secure the early publication of an acceptable article in the first issue to go to press after its acceptance by arranging for the article to be printed as extra pages, increasing the number of pages that subscribers receive. Information about charges for early publication may be obtained from the Managing Editor.

National Auxiliary Publications Service. To save printing costs, materials supplementary to articles may be deposited with the NAPS. To obtain copies, order by Document Number, as cited in the reference footnote, from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001. The indicated fee must be remitted in advance.

Abstracts. Manuscripts must be accompanied by an abstract of 100–120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or they may be obtained from the Editor or from the APA Central Office.

Copyrighted material. The following policies govern reprinting or other forms of reproduction of APA-copyrighted journal materials. Permission by the American Psychological Association must be requested for reprinting all tables, figures, and text of more than 500 words (total from one article) in length. There will be a charge of \$5.00 per page for such permission except when the requester is the author of the article to be reprinted, or the article is to be reproduced in a limited number of copies solely for instructional purposes.

Reprints. Authors receive 50 reprints gratis. Additional quantities may be ordered when returning proofs.

FIVE ATTEMPTS TO REPLICATE THE EXPERIMENTER BIAS EFFECT¹

THEODORE XENOPHON BARBER²

Medfield Foundation, Herding, Massachusetts

ALBERT FORGIONE

Boston University

JOHN F. CHAVES

Massachusetts Bay Community College

DAVID S. CALVERLEY

Worcester Junior College

JOHN D. McPEAKE

Fisher Junior College

AND BARBARA BOWEN

Regis College

Five investigations, involving 501 Ss and 51 Es, were conducted to cross-validate an earlier study (Rosenthal & Fode, 1963, Exp. 1) that had clearly demonstrated the Experimenter Bias Effect. Each of the five investigations failed to demonstrate that Es' expectancy-biases influence their results (overall $F < 1.0$). The five investigations are related to other studies in this area, the majority of which also failed to demonstrate the Experimenter Bias Effect. It is concluded that the effect is more difficult to demonstrate than was implied in several recent reviews and that, at the present time, it is not known what preconditions are necessary to obtain it.

A recent text (Rosenthal, 1966) delineated several ways that an experimenter might inadvertently influence the results of his research. These "experimenter effects" include an Experimenter Personal-Attributes Effect in which the characteristics or personality traits of the experimenter—for example, his sex, age, race, anxiety, need for approval, hostility, and dominance—influence his Ss' performance. They also include an Experimenter Bias Effect in which E's expectancies, hypotheses, desires, or biases affect the results. For instance, an Experimenter Bias Effect would be present if Es obtain significantly higher scores from Ss on a standard test when they expect high rather than low scores.

This paper is concerned only with the Experimenter Bias Effect. There is little doubt that Es generally expect or desire to obtain different results from Ss assigned to different

experimental treatments and that they rarely exclude the possibility that their Ss' responses are influenced by these expectancies and desires. If the Experimenter Bias Effect is pervasive—if Es inadvertently influence their Ss to respond in such a way as to confirm their expectancies or desires—a substantial proportion of the "facts" of present-day psychology would be open to question and many if not most earlier psychological investigations would need to be rerun to determine if the results were due to Es' biases. Furthermore, as Rosenthal (1966, pp. 331-400) has emphasized, if the Experimenter Bias Effect has general validity, it would be necessary in further research either (a) to institute complex procedures to obviate the effects of the experimenters' expectancies and desires or (b) to remove human beings from the conduct of experiments—replacing them by automated procedures.

In all of the formal experimental studies that aimed to demonstrate the Experimenter Bias Effect, the Es were college students (Rosenthal, 1966). The criterion instruments in most of these investigations was a person-perception task in which S was shown a series of photographed faces by the student-E and was asked to rate on a scale whether he judged that the persons depicted had been "experiencing failure" or had been "experienc-

¹ These investigations were supported by research grants (MH-07003 and MH-11521) from the National Institute of Mental Health, United States Public Health Service, to T. X. Barber.

² The authors are grateful to Robert Rosenthal for supplying the photographs for the person-perception task and to the following for criticizing the manuscript: Robert Rosenthal, Maurice J. Silver, Ralph L. Rosnow, Herbert J. Greenwald, Nicholas Spanos, and Dick Murray. Requests for reprints should be sent to Theodore X. Barber, Medfield State Hospital, Medfield, Massachusetts 02042.

ing success." The rating scale employed ran from -10 (extreme failure) to +10 (extreme success) with intermediate labeled points. In investigations which clearly showed the Experimenter Bias Effect (e.g., Rosenthal & Fode, 1963, Exp. 1), one group of student-*Es* was told, prior to the experiment, that their *Ss* were expected to give high ratings (averaging +5) on the person-perception task and another group of student-*Es* was told that their *Ss* would give low ratings (averaging -5). In addition, all student-*Es* were told that (a) the expected results had been "well established" in previous investigations, (b) they would conduct the experiment to obtain practice in "duplicating" the previously established findings, (c) if their results came out "properly—as expected" they would be paid \$2.00 per hour, and (d) if their results did not come out "properly" they would be paid \$1.00 per hour (Rosenthal & Fode, 1963, p. 507). An important consideration is relevant here: The possibility was not excluded in these investigations that, in order to obtain more money or to make the results come out "properly—as expected," some or many of the student-*Es* may have intentionally failed to conduct the experiment in the way they were told to carry it out. For instance, the student-*Es* may have (a) verbally reinforced their *Ss* ("good," "fine," "that's excellent") when they gave the "proper" ratings or (b) misreported their *Ss'* ratings.

Our first investigation was conducted with the above considerations in mind. When planning the first investigation, the project-coordinator (TXB) hypothesized that the student-*Es* would obtain high or low ratings from their *Ss* in line with their expectancies. The main purpose of the first investigation, however, was not simply to demonstrate the Experimenter Bias Effect but to determine whether the effect was due to the student-*Es* verbally reinforcing their *Ss* for "proper" ratings or misreporting their *Ss'* ratings. To ascertain how the student-*Es* tested their *Ss*, we carefully concealed a tape recorder in the experimental room and we did not inform either the *Es* or the *Ss* that the recorder was present. However, when we (TXB and DSC) analyzed the results of the first investigation, we were surprised to find that the Experi-

menter Bias Effect was not manifested; student-*Es* given an expectancy for high (+5) ratings, those given an expectancy for low (-5) ratings, and a control group of student-*Es* not given expectancies did not differ in the ratings they obtained ($F < 1.0$).

Even though the Experimenter Bias Effect was not demonstrated in our first investigation, we proceeded to listen to the tapes to determine how the student-*Es* had conducted the experiment. Unfortunately, the tape recorder had been placed at too great a distance from the *Es* and *Ss* and, since the tapes were barely audible, we were unable to analyze them. However, our inability to analyze the verbal interactions between the student-*Es* and the *Ss* did not alter the major finding of the first investigation, namely, that the Experimenter Bias Effect was not noticeably present.

Having failed to demonstrate the Experimenter Bias Effect in our first investigation, we (TXB and DSC) then asked: Is this effect replicable? To answer this question, we set up four additional investigations. The principal investigators who were asked to supervise the second, third, fourth, and fifth investigations (AF, JDM, JFC, and BB) were not informed of our previous negative results. Investigations 2-5 were conducted in the same way as Investigation 1 with the exception that tape recorders were not concealed in the experimental rooms.

We attempted to design each of the five investigations so as to replicate, as closely as possible, the design of a previous study (Rosenthal & Fode, 1963, Exp. 1) that had clearly demonstrated the Experimenter Bias Effect. Our investigations, however, extended the Rosenthal and Fode study in that they included an additional (control) group of experimenters who were not given expectancies. The method and procedures in each of our investigations were as follows.

METHOD

A total of 558 individuals participated in the five investigations: 1 project coordinator (TXB), 5 principal investigators, 51 student-*Es*, and 501 *Ss*.

Principal Investigators

Four of the five principal investigators (DSC, AF, JDM, and JFC) were teaching classes in introduc-

tory psychology at four college in Massachusetts. The fifth principal investigator (BB) was a third-year honors student majoring in sociology at a fifth college.

After each principal investigator agreed to supervise an investigation at his college he was told by the project coordinator that he was to (a) use a table of random numbers to select about nine experimenters and about 90 Ss from the students available to him, (b) randomly administer one of three sets of expectancy instructions (+5, -5, or control expectancy instructions) to each of his student-Es, (c) administer the person-perception task to each student-E as a "warm-up," and (d) arrange for each student-E to test about 10 randomly selected Ss on the person-perception task. Finally, each principal investigator was asked to state his own expectancies concerning the outcome of his investigation.³

Experimenters

Four of the five principal investigators randomly selected the Es from the students in their introductory psychology classes; the fifth principal-investigator (BB) selected the Es at random from the list of freshmen and sophomores at her college. Investigation 4 included 15 student experimenters and the remaining four investigations each included nine student-Es.

Subjects

The Ss were freshmen and sophomores at the five colleges enrolled in introductory psychology courses. The number of Ss participating in Investigations 1 to 5, respectively, were 81, 90, 90, 150, and 90.

Procedure

There were two parts to the procedure. In Part 1 the principal investigators instructed the student-Es and in Part 2 the student-Es tested the Ss.

Part 1. Each principal investigator read the instructions that are presented verbatim by Rosenthal and Fode (1963, Appendix C) to those student-Es who

³One of the five principal-investigators (BB) apparently had not heard about previous research concerned with the Experimenter Bias Effect and did not proffer any expectancies concerning the outcome of the investigation. The other four principal investigators, who were teaching courses in psychology, had previously heard about the research in experimenter bias. Three of these (DSC, JDM, and JFC) predicted that the ratings would be highest under the +5 expectancy condition, next highest under the control expectancy condition, and lowest under the -5 expectancy condition. The remaining principal investigator (AF) predicted that the student Es "would lean over backwards not to bias their subjects" and, consequently, student-Es expecting high ratings would obtain significantly lower ratings than those expecting low ratings. As will be seen when we present the results, in no case were the expectancies of the principal investigators confirmed.

had been randomly assigned to the high (+5) expectancy condition and to the low (-5) expectancy condition. The instructions read to the student-Es who were randomly assigned to the control expectancy condition differed in that nothing was said concerning "expected" or "proper" results. After the principal investigator read the (high, low, or control expectancy) instructions to each student-E, he asked each student-E to read the instructions to himself.

In accordance with the procedures used by Rosenthal and Fode (1963, Exp. 1), each student-E was next asked, as a "warm-up," to rate the standardized set of photographs himself.

Part 2. The heart of the experiment came next, when each student-E administered the person-perception task to his Ss. The procedures for administering the task were identical with those of the earlier study (Rosenthal & Fode, 1963, Appendix B).⁴

RESULTS

Table 1 presents, separately for each of the five investigations, the mean ratings per photograph obtained by each student-E from his Ss, the mean ratings obtained by student-Es assigned to the high (+5), control, and low (-5) expectancy conditions, and the *F* ratios yielded by one-way analyses of variance applied to the latter means. The final row of Table 1 presents the overall mean ratings obtained under the high, control, and low expectancy conditions in the five investigations combined. These results can be briefly summarized as follows:

1. One-way analyses of variance performed separately on the results of each of the five investigations yielded *F* ratios for expectancy conditions which were in each instance either less than 1.0 or close to 1.0. These *F*s, of course, are far from significant. A two-way analysis of variance (Expectancy Conditions \times Replications) was performed upon the overall results of the five investigations (Winer, 1962, pp. 241-244). This analysis (summarized in part in the final note of Table 1) showed a far from significant effect for expectancy conditions ($F < 1.0$), a nonsignificant effect for replications ($p = .18$), and

⁴In these investigations Es and Ss were undergraduates taking their first college course in psychology. Their texts and lectures in psychology had not referred to the Experimenter Bias Effect. In none of the five investigations was there any indication that the student-Es were aware that they themselves were the objects of investigation or that prior awareness of the experimenter bias phenomenon on the part of the student-Es or Ss influenced the results.

TABLE 1

MEAN RATINGS ON PERSON PERCEPTION TASK UNDER HIGH, CONTROL, AND LOW EXPECTANCY CONDITIONS

Investigation	High expectancy	Control expectancy	Low expectancy
1. Worcester Junior College ^a	-.63 +.16 +.99 +.17	+.14 -.03 -.39 -.09	+.17 -.17 +.87 +.29
<i>M</i>			
2. Boston University ^b	+.12 -.26 +.34 +.07	+.79 +.70 -.29 +.40	-.29 +.12 -.10 -.09
<i>M</i>			
3. Fisher Junior College ^a	+.67 +.27 +.72 +.55	+.66 +.52 -.02 +.39	+.34 +.80 +.73 +.62
<i>M</i>			
4. Massachusetts Bay Community College ^c	+.66 -.24 +.87 -.64 -.10	.00 -.28 +.14 .00 -.33	-.02 +.72 -.37 -.61 +.22
<i>M</i>			
5. Regis College ^d	+.24 +.28 +1.22 +.58 +.27	-.18 -.25 -.02 -.15 +.07	+1.20 +.10 -.06 +.41 +.21
<i>M</i>			
Overall mean of five investigations ^e			

Note.—In Investigation 1 each cell mean is based on nine Ss. In Investigations 2-5 each cell mean is based on 10 Ss.
^a $P < 1.0$; $df = 2/6$.
^b $P < 1.14$; $df = 2/6$.
^c $P < 1.0$; $df = 2/12$.
^d $P < 1.66$; $df = 2/6$; $p > .25$.
^e $P < 1.0$; $df = 2/36$ for expectancy conditions.

a far from significant interaction ($F < 1.0$). In brief, analyses of variance performed separately for each of the five investigations and for all five investigations taken in combination consistently failed to show that *Es* with different expectancy-biases obtained different ratings from their *Ss*. Nonparametric Kruskal-Wallis tests yielded the same null results.

2. None of the five investigations showed a trend in line with the Experimenter Bias Effect; that is, no tendency was noticeable in any one of the five investigations for the mean ratings to be slightly higher under the +5 expectancy condition, next highest under the control expectancy condition, and lowest under the -5 expectancy condition. In Investigations 1 and 3 the highest mean rating (in absolute terms) was obtained under the low

(-5) expectancy condition. In Investigation 2 the highest mean rating was obtained under the control expectancy condition rather than under the high (+5) expectancy condition. In Investigations 4 and 5 the lowest mean rating was obtained under the control expectancy condition rather than under the low expectancy condition. Taking all five investigations in combination, the overall mean ratings (.27, .07, and .21 under the high, control, and low expectancy conditions, respectively) were not consistently in the expected direction—the mean under the control condition is lower than the mean under the low expectancy condition.

In brief, these five investigations, taken separately and also in combination, failed to demonstrate the Experimenter Bias Effect.⁵

DISCUSSION

Each of our five investigations attempted to replicate the procedures used in an earlier experiment (Rosenthal & Fode, 1963, Exp. 1) that had clearly demonstrated the Exper-

⁵ We gave the raw data of these investigations to Robert Rosenthal for further analyses. He (Rosenthal, 1967b) informed us that if the data are partitioned in various ways and chi-square tests are performed, some of the chi-squares yield *ps* ranging from .16 to .10. However, even if we accept *ps* of .16 and .10 as significant and even if we accept the results of the chi-square tests and disregard the results of the Kruskal-Wallis tests and of the analyses of variance (which yielded an overall *F* smaller than 1.0 with an associated *p* that was very far from significant), we still cannot conclude that our data support the experimenter bias hypothesis because the means and the medians were not in the direction predicted by this hypothesis—the control (no-expectancy) *Es* obtained lower mean and median ratings than the *Es* who expected low ratings.

Rosenthal (1967b) also informed us that significant results were obtained in five statistical tests (chi-squares and a binomial test) which were performed on the data obtained from those *Ss* who gave ratings of 2.0 or higher on the person-perception task. We cannot accept these tests as showing the Experimenter Bias Effect because only a very small percentage of *Ss* (7% to 11%), who scored 2.0 or above, were arbitrarily selected for these analyses; if equally arbitrary statistical tests had been performed on the data of *Ss* scoring 3.0 or higher, or below 2.0, the results would have been far from significant. The issue here, pertaining to the misleading conclusions that derive from arbitrary selection of data for postmortem analysis, are discussed in detail elsewhere (Barber & Silver, 1968b).

menter Bias Effect. Although we attempted to follow the procedures of the earlier study, we failed to demonstrate the effect. The results of our investigations appear to indicate that the Experimenter Bias Effect may be more difficult to demonstrate than was implied in recent reviews (Kintz, Delprato, Mettee, Persons, & Schappe, 1965; Rosenthal, 1963, 1964, 1966, 1967a). This conclusion is supported by the following considerations.

When we (TXB and DSC) began the five investigations, we had received the impression from reading the review papers in this area that the Experimenter Bias Effect had been demonstrated in practically all studies that had been designed to show it. In fact, we *assumed* that the effect would be demonstrated in the first investigation and we aimed to ascertain how the effect would be mediated. After failing to show the effect in five investigations, we (Barber & Silver, 1968a) looked very closely at all of the extant investigations that had attempted to demonstrate that the expectancy-biases of student-*Es* influence their *Ss'* responses. Contrary to our original assumption, we found that the majority of studies in this area had *failed* to show the Experimenter Bias Effect (Barber & Silver, 1968a). After considering the matter further, we were able to delineate several interrelated reasons why we and other investigators had failed to demonstrate the effect. To show that expectancies of student-*Es* significantly bias their *Ss'* responses, it is necessary for expectancies to be transmitted from the principal investigator to the student-*E* and from the student-*E* to *S*. In fact, it appears that a complex eight-step transmission process is involved (cf. McGuire, 1967):

1. The student-*E* must attend to the expectancy communication from the principal investigator.

2. The student-*E* must comprehend the expectancy communication.

3. The student-*E* must retain the communication.

4. The student-*E* (intentionally or unintentionally) must attempt to transmit the expectancy to *S*.

5. The *S* ("consciously" or "unconsciously") must attend to the expectancy communication from *E*.

6. The *S* ("consciously" or "unconsciously") must comprehend *E's* expectancy.

7. The *S* ("consciously" or "unconsciously") must retain *E's* expectancy.

8. The *S* (wittingly or unwittingly) must act upon (give responses in harmony with) *E's* expectancy.

In investigations that failed to demonstrate an Experimenter Bias Effect, the transmission process could have broken down at any of these eight links in the chain. There are many possibilities here: the student-*E* may have failed to comprehend or to retain the expectancy communication from the principal investigator; the student-*E* may have comprehended and retained the expectancy communication but may have failed to transmit it to *S*; *S* may have comprehended *E's* expectancy but may have failed to respond in the way expected; etc. It is apparent that systematic research is needed to determine under what circumstances the transmission of expectancies is most likely and also least likely to break down at each of the eight steps in the chain.

Although we and other investigators have failed to demonstrate the Experimenter Bias Effect, it must be emphasized that the effect has been demonstrated in a number of studies (Barber & Silver, 1968a). At present it is unclear why some studies succeeded and others failed in this demonstration. However, Rosenthal (1967b) has pointed out to us that the contradictory results obtained in our investigations and in the Rosenthal and Fode (1963, Exp. 1) study may be due to subtle differences such as the following: in our investigations *Es* and *Ss* were enrolled in the same introductory psychology course; in contrast, in the Rosenthal and Fode study *Es* were enrolled in an undergraduate experimental psychology laboratory whereas *Ss* were enrolled in an introductory psychology course and three of the 10 *Es* were first-year graduate students. These subtle factors are related to the following differences between the two sets of studies that may have produced the contradictory results: (a) in our investigations, but not in the Rosenthal and Fode study, *Es* were previously acquainted with almost all of their *Ss*; (b) in our investigations *Es* and *Ss* were of equal status whereas in the Rosenthal and Fode study *Es* were of slightly

higher status, and (c) in the Rosenthal and Fode study *Es* were in a situation (an undergraduate experimental psychology laboratory) in which "widespread data fabrication" is common (Rosenthal, 1966, p. 35). Whether or not these factors are sufficient to account for the contradictory results of our investigations and of the Rosenthal and Fode study is not clear at the present time. Rosenthal (1966, pp. 239-241) has presented data which suggest the hypothesis that the Experimenter Bias Effect is more readily obtained when male *Es* are acquainted with their *Ss* and when female *Es* are unacquainted with their *Ss*. With respect to the status of *Es* vis-à-vis *Ss*, the data are also contradictory (Rosenthal, 1966, pp. 241-245): One study (Vikan-Kline, 1962) indicated that high-status rather than low-status *Es* are more likely to bias their results whereas another study (Laszlo & Rosenthal, 1967) indicated that low-status rather than high-status *Es* are more likely to bias. Finally, although student-*Es* who are enrolled in an undergraduate experimental psychology laboratory may be prone to fudge their data, we have no direct evidence that the presence of this factor accounted for the positive results in the Rosenthal and Fode study or that its absence accounted for the negative results in our investigations.

In conclusion, it appears that the Experimenter Bias Effect is more difficult to demonstrate than was implied in several recent reviews (Rosenthal, 1963, 1964, 1966, 1967a) and further research is needed to determine exactly what preconditions are necessary to obtain the effect.

REFERENCES

- BARBER, T. X., & SILVER, M. J. Fact, fiction, and the Experimenter Bias Effect. *Psychological Bulletin*, 1968, 70(6, Pt. 2), 1-29. (a)
- BARBER, T. X., & SILVER, M. J. Pitfalls in data analysis and interpretation: A reply to Rosenthal. *Psychological Bulletin*, 1968, 70(6, Pt. 2), 48-62. (b)
- KLINTZ, B. L., DELPRATO, D. J., METTIE, D. R., PERSSONS, C. E., & SCHAPPE, R. H. The experimenter effect. *Psychological Bulletin*, 1965, 63, 223-232.
- LASZLO, J. P., & ROSENTHAL, R. Subject dogmatism, experimenter status and experimenter expectancy effects. Department of Social Relations, Harvard University, 1967. (Mimeo)
- MCGUIRE, W. J. Personality and susceptibility to social influence. In E. F. Borgatta & W. W. Lambert (Eds.), *Handbook of personality theory and research*. Chicago: Rand McNally, 1967.
- ROSENTHAL, R. On the social psychology of the psychological experiment: The experimenter's hypothesis as unintended determinant of experimental results. *American Scientist*, 1963, 51, 268-283.
- ROSENTHAL, R. The effect of the experimenter on the results of psychological research. In B. A. Maher (Ed.), *Progress in experimental personality research*. Vol. 1. New York: Academic Press, 1964.
- ROSENTHAL, R. *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts, 1966.
- ROSENTHAL, R. Covert communication in the psychological experiment. *Psychological Bulletin*, 1967, 67, 356-367. (a)
- ROSENTHAL, R. The eternal triangle: Investigators, data and the hypothesis called null. Unpublished manuscript, Harvard University, 1967. (b)
- ROSENTHAL, R., & FODE, K. Three experiments in experimenter bias. *Psychological Reports*, 1963, 12, 491-511.
- VIKAN-KLINE, L. L. The effect of an experimenter's perceived status on the mediation of experimenter bias. Unpublished master's thesis, University of North Dakota, 1962.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

(Received January 2, 1968)

Editor's Note

The following three discussions were prepared at the invitation of the Editor because of the earlier article's challenge to the timely and significant work reported elsewhere by Robert Rosenthal.

ON NOT SO REPLICATED EXPERIMENTS AND NOT SO NULL RESULTS¹

ROBERT ROSENTHAL²

Harvard University

Recent research by Barber, Calverley, Forgione, McPeake, Chaves, and Bowen (1969) has been interpreted as a failure to replicate an experiment (Rosenthal & Fode, 1963) showing significant effects on *Ss*' responses of *Es*' expectancies. The purpose of this paper is to show (a) that the above research, while helpful in its own right, can in no sense be regarded a serious effort to replicate, (b) that, in any case, the integrity of the null hypothesis cannot be staunchly defended by an overall low F ($N = 51$ *Es*) in the face of an overall χ^2 ($N = 501$ *Ss*) with an associated p value of .007 and (c) that science is not often served best by employing the null-hypothesis decision procedure in such a way as to avoid looking closely at data.

This paper constitutes a comment on and an addendum to the research by Barber, Calverley, Forgione, McPeake, Chaves, and Bowen (1969). It was the intent of these investigators to replicate the first of a long series of studies of the effects of the expectancy of *E* on his *Ss*' responses. In a series of five samples, one-third of the *Es* were led to expect from their *Ss* positive ratings of the success of persons pictured in photographs. Another third of the *Es* were led to expect negative ratings, and one-third of the *Es* were given no expectation. This last group was a very instructive addition and, as we shall see when we look more closely at the data, provides us with important new evidence. First, however, we must consider the proposition that useful though it was, the research described by Barber et al. can by no stretch of the imagination be regarded as a serious effort at replication in the usual sense.

PROCEDURAL DIFFERENCES

Some more general issues surrounding the problem of replication in the behavioral sciences have been raised elsewhere (Rosenthal, 1966) and need not be repeated here. It is enough to say that psychological research is characterized not only by too few efforts to replicate but also by a very high proportion of failures to replicate. In most replication efforts *Ss* are different, often *E* is different, the time of year, the locale, and the physical scene are different. With all these built-in differences, careful replication requires at least that the procedures of the original research be

followed closely. In the following ways that was not done in the research by Barber et al.

Type of subject. The original experiment employed, as *Ss*, students enrolled in a coeducational state university. Only 18% of the *Ss* of the "replication" were from a similar population. Another 18% were from a 4-year Catholic college for girls and the remaining 64% were from junior colleges. Ordinarily such sampling differences might be regarded as trivial. However, because of the casual method of sampling *Es*, to be discussed presently, these differences become quite critical.

Subjects per experimenter. In the original experiment each *E* contacted an average of over 20 *Ss*. In the "replication" each *E* contacted an average of less than 10. Such a difference ought not to affect the magnitude of the grand means but it should affect the reliability of the mean response obtained by each *E* and it was these mean responses that were used as the basic data by Barber et al.

Sex of experimenter. In the original experiment most (80%) of the *Es* were males while in the "replicate" most (59%) of the experimenters were females. For some studies this difference might not be important but evidence has been presented (Rosenthal, 1966) that female *Es*, especially when contacting male *Ss*, are less likely to show expectancy effects. Indeed, they have been shown sometimes to obtain data significantly opposite to that expected.

Experimenter status. In the original study *Es* were at least advanced undergraduate majors in psychology, 30% were graduate students and all had laboratory course experience. The *Ss* in the original, as in the "replicate," were beginning students. In all our research involving several hundred *Es* and thousands of *Ss* we have always been careful to preserve the differential in status and educational background between *Es* and *Ss*, and

¹Preparation of this paper was facilitated by Grant GS-1741 from the Division of Social Sciences of the National Science Foundation.

²Requests for reprints should be sent to Robert Rosenthal, Department of Social Relations, Harvard University, Cambridge, Massachusetts 02138.

that fact has been explicitly emphasized (Rosenthal, 1966, p. 306). Not only the evidence from our own research program but the weight of the evidence from the general literature on social influence suggests that higher status influencers are more likely to be successful (Rosenthal, 1966).

In the present study *Es* were selected at random from the subject population. These *Es*, then, were classmates of their *Ss*, with not even the slight advantage of knowing something more about laboratories, psychology, research, or science than their *Ss*. Perhaps most important of all is the fact that their *Ss* knew this. It is possible that when an *E* is of very high status he can influence unintentionally an *S* of equally high status. There are no data on this point but high absolute status may be sufficient without a status differential. In the case of *Es* employed by Barber et al., not only was there no status differential between *E* and *S* but the absolute status of *E* was about as low as it could get for a collegiate population. Most of the *Es* were beginning junior-college students.

If the research by Barber et al. had shown no effects of *E* expectancy (a conclusion drawn by the authors but as we shall see, one that is not all so clearly warranted by the data) an earlier hypothesis (Rosenthal, 1966) might have been given further indirect support. That hypothesis, and there are data to support it, is that more competent, more professional, more experienced *Es* of higher status are more likely unintentionally to affect their *Ss'* responses by virtue of their expectancy.

TWO TYPES OF REPLICATION

It may be useful at this point to introduce a terminological and conceptual distinction between two types of replication.

Replications of tenability. Replication of the first kind, or replication of an experiment, is characterized by an attempt to reproduce as closely as possible all of the conditions and procedures of an earlier experiment. It is a type of replication addressed to the specific tenability of a specific outcome of a specific experiment. Such replications help us to rule out the operation of unsuspected idiosyncrasies in the original experiment, idiosyncrasies that made the original result unlikely to recur. Such idiosyncrasies include such human factors as errors of observation, recording, computation, and report. They include also such nonhuman factors as undetected equipment failures and differences in the physical arrangements of the laboratory. Finally, they include also such individual difference variables as the particular *E's* characteristics, the particular

sample of *Ss* standing in a given relationship to some important national or local event, etc.

Quite clearly, it was this type of replication that Barber et al. intended to conduct. Thus, they wrote of their attempt "... to replicate, as closely as possible ..." It is only in the sense in which they intended it, the sense of a replication of tenability, that their "replication" cannot be considered even a close approximation. That is not to say that their research is addressed to the wrong question by virtue of their altered procedures, but only that it is addressed to a different one.

Replications of generality. Replication of the second kind, or replication of a relationship, is characterized by a purposeful effort to change some potentially important aspect of a previously conducted experiment. It is a type of replication that usually occurs after some replication of tenability has been conducted and it is addressed to a testing of the limits of the domain over which the earlier found relationship obtains. It would seem quite reasonable to view the research by Barber et al. as a replication of this second type.

Similarly, recent summaries of experiments showing the effects of one person's expectations about the behavior of another person are essentially summaries of replications of generality. One of these summaries is a reconsideration of over 30 studies considered by Barber and Silver (1968)^a in which the relevant interpersonal expectation is held by an *E* and is about his *Ss'* behavior (Rosenthal, in press). Another summary is of studies in which the relevant interpersonal expectation is held by a teacher and is about her pupils' intellectual growth (Rosenthal & Jacobson, 1968). Different as these two types of studies are, they can nevertheless be regarded as replications of a relationship though not, of course, as replications of specific laboratory experiments.

ON LOOKING MORE CLOSELY AT DATA

On the basis of their analysis of the data, Barber et al. concluded that they had failed to demonstrate the effect of *Es'* expectancies. If that were clearly the case we should not be unduly surprised. Barber et al. give many good reasons in their paper why such "null" results might have been obtained. Indeed there have been a number of studies which quite clearly showed no effects

^a Barber et al. cite the conclusion by Barber and Silver (1968) that the majority of the studies they selected for review failed to show effects of *Es'* expectancies. It would take us too far afield to discuss that conclusion here, but it has been discussed elsewhere in some detail (Rosenthal, in press).

TABLE 1

MEAN RATINGS OF PHOTOGRAPHS BY
501 SUBJECTS

Photo ratings	Expectancy of experimenter			
	+5	None	-5	Σ
+3.00 and above	9	5	6	20
+2.50 to +2.99	9	6	0	15
+2.00 to +2.49	8	2	5	15
+1.50 to +1.99	11	19	10	40
+1.00 to +1.49	14	15	25	54
+0.50 to +0.99	14	17	28	59
0.00 to +0.49	30	20	29	79
-0.50 to -0.01	21	26	21	68
-1.00 to -0.51	18	22	14	54
-1.50 to -1.01	16	11	12	39
-2.00 to -1.51	4	10	8	22
-2.50 to -2.01	5	9	0	14
-2.51 and below	8	5	9	22
Σ	167	167	167	501*

* $\chi^2 = 44.93$, $df = 24$, $p < .007$, two-tailed test.

of E expectancy. It would be an unusual area of behavioral research that turned up only p values of .05 or less. Our purpose here is only to give the reader some additional data to guide his judgment.

There were 51 E s and 501 S s employed in the research reported by Barber et al. They presented the results based on the 51 E s. Table 1 shows the results based on the 501 S s. For this single overall analysis of all S s' photo ratings, intervals of .50 were employed, any smaller interval being such as to reduce the expected frequencies below the rounded value of 5. The χ^2 based on Table 1 showed an associated two-tailed p less than .007, but a χ^2 based on 13 intervals in three conditions is not so directly informative even when it seems clear that the entries are not randomly arrayed. We shall first want to see which of the three groups show large differences from one another.⁴

The two groups in which we are most interested, since they were most like the groups employed in other such studies, are those in which E s were given positive (+5) and negative (-5) expectations. That χ^2 was 25.59, $df = 12$, $p < .02$, two-tailed. Inspection of Table 1 suggests that this result was due to the E s' expecting positive photo ratings obtaining too many of the very high ratings but too few of the slightly high ratings. When we compare the ratings made by S s contacted by E s expecting negative photo ratings with the ratings made by S s contacted by the E s given no expectancy we find a χ^2 of 29.72, $df = 12$, $p < .004$, two-tailed. Relative to E s given no expectations, E s led to expect low photo ratings obtained too few very high, too many slightly high, and too few negative ratings. The comparison of photo ratings given by S s whose E s had been led to expect high photo ratings with responses of S s of the control condition showed only a much smaller difference ($\chi^2 = 16.05$, $df = 12$, $p < .20$, two-tailed). The direction of differences was such that E s expecting high photo ratings, relative to the controls, obtained too many very high ratings and too few slightly high ratings.

The overall pattern of results described, while

⁴ I want to thank T. X. Barber for making these data available. In an earlier communication to Barber et al. and cited by them, the overall χ^2 was less significant because coarser grouping had been employed (below -2.00 to above +2.00 advancing in intervals of 1.00). In that communication, a number of additional subanalyses were reported which were interpreted by Barber et al. as alternative analyses rather than simply as an attempt to check the consistency from sample to sample of the overall effect obtained.

surprising and certainly quite unpredicted, nevertheless does not make for a strong defense of the null hypothesis that E s' expectations have no effect on their S s' ratings of photographs. The curvilinear nature of the effect obtained does help us to understand, however, why the analysis of variance, with its insensitivity to such curvilinearity, would show no significant effect of E expectations.

How might we account for the effects obtained? If sheer speculation were permitted, we might wonder whether our sample was made up of two types of S s. One type may have behaved as did the S s of the earlier research in which S s responded substantially in the direction of their E 's expectation. The other type of S may have rebelled slightly, giving E responses in a direction slightly opposite to that expected by E . On this point we must agree with Barber et al.; more research is needed.

SCIENCE AND STATISTICAL PROCEDURES

Barber et al. conducted a series of studies on E expectancy effects and performed overall analyses of variance which led to F s too small for them to reject the null hypothesis. In spite of certain fairly apparent trends in the data which were pointed out to the authors, they felt it was statistically unjustifiable to look more closely at their data. But that closer look shows that much was missed by their procedure though theirs appears to be a defensible procedure and, in the eyes of some psychologists, the only "proper"

one. The null-hypothesis decision procedure has been discussed in detail elsewhere (Bakan, 1966; Rozeboom, 1960). The view of science that underlies many workers' usage of this procedure seems to have the effect of protecting the investigator from a full view of his data. The null-hypothesis decision procedure is perhaps too often employed as an opaque screen placed over the potential for learning something new.

The position of the present writer is that it is not from data that investigators require protection. Some protection is required from prematurely firm conclusions drawn from a set of data, but the protection of a low alpha is not a very secure safeguard. The only ultimate protection is replication. It is often said that what is found serendipitously in the data requires replication. That seems unquestionably true. The implied corollary, however, that a predicted finding requires less replication does not seem to follow. All findings require replication, those that are clear and predicted no less than those that are less clear and "only" discovered. Even this cannot be interpreted too narrowly, however. If the predictions are based on earlier empirical tests of the same prediction we can have increasing confidence as more supporting evidence is obtained (Rosenthal, in press).

There is a frequently held, and implicitly so, assumption that relationships that are not immediately obvious from the way in which an investigator happens to have arrayed his data are somehow less real. It is as though a close look at the data changed the data and made falsely "true"

what was truly "false." The analogy to such an assumption is the case of the lookout in the crow's nest who samples the horizon for a view of land. If, deviating from the visual sweep prescribed by his captain, he should see some unexpected land, shall we doubt the lookout's vision because the captain's orders did not include the view?

REFERENCES

- BAKAN, D. The test of significance in psychological research. *Psychological Bulletin*, 1966, 66, 423-437.
- BARBER, T. X., CALVERLEY, D. S., FORGIONE, A., McPEAKE, J. D., CHAVES, J. F., & BOWEN, B. Five attempts to replicate the experimenter bias effect. *Journal of Consulting and Clinical Psychology*, 1969, 33, 1-6.
- BARBER, T. X., & SILVER, M. J. Fact, fiction, and the experimenter bias effect. *Psychological Bulletin*, 1968, 70(6, Pt. 2) 1-29.
- ROSENTHAL, R. *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts, 1966.
- ROSENTHAL, R. Experimenter expectancy and the reassuring nature of the null hypothesis decision procedure. *Psychological Bulletin*, in press.
- ROSENTHAL, R., & FODE, K. L. Psychology of the scientist: V. Three experiments in experimenter bias. *Psychological Reports*, 1963, 12, 491-511.
- ROSENTHAL, R., & JACOBSON, L. *Pygmalion in the classroom*. New York: Holt, Rinehart & Winston, 1968.
- ROZEBOOM, W. W. The fallacy of the null-hypothesis significance test. *Psychological Bulletin*, 1960, 57, 416-428.

(Received June 17, 1968)

INVALID ARGUMENTS, POSTMORTEM ANALYSES, AND THE EXPERIMENTER BIAS EFFECT¹

THEODORE XENOPHON BARBER²

Medfield Foundation, Harding, Massachusetts

Rosenthal (1969) presented four arguments aiming to show that the five experiments by Barber, Calverley, Forgione, McPeake, Chaves, and Bowen (1969) cannot be viewed as serious efforts to replicate the Rosenthal-Fode (1963) study. These arguments are shown to be invalid. Rosenthal's critique also presented a postmortem analysis of the Barber et al. data. A further look at Rosenthal's analysis indicates that the Barber et al. experiments cannot be interpreted as confirming the experimenter bias hypothesis that *Es* obtain results in line with their expectancies. It is concluded that the Experimenter Bias Effect is very difficult to demonstrate.

Rosenthal's (1969) critique of the five experiments by Barber, Calverley, Forgione, McPeake, Chaves, and Bowen (1969) makes two points: (a) The Barber et al. experiments did not closely replicate the procedures used in the original study by Rosenthal and Fode (1963, Exp. 1). (b) Although Barber et al. concluded that their data did not show an Experimenter Bias Effect, a different way of analyzing their data yields a significant chi-square. These two points will be discussed in turn.

DISCREPANCIES IN REPLICATING THE ROSENTHAL- FODE PROCEDURES

Rosenthal's critique lists four discrepancies between the Barber et al. experiments and the original study, by Rosenthal and Fode, that was to be replicated:

Type of subjects. The Rosenthal-Fode study used *Ss* from a coeducational state university whereas the five Barber et al. experiments were conducted in a coeducational nonstate university, a Catholic college for girls, and three junior colleges. Rosenthal's critique states that "because of the casual method of sampling *Es* . . . these differences become quite critical." The present writer fails to see why these differences are "critical." If the Experimenter Bias Effect has any degree of generality it should not be limited to coeducational state universities. If Rosenthal is seriously contending that the Experimenter Bias Effect is more readily obtained in coeducational state universities than in other types of

colleges or universities, he should present data to support the contention.

Subjects per experimenter. In the Rosenthal-Fode study, each *E* contacted an average of 20 *Ss* whereas in the Barber et al. experiments each *E* contacted about 10 *Ss*. The point made in Rosenthal's critique, that these differences would affect "the reliability of the mean response obtained by each *E*," seems to imply that the Barber et al. experiments were less sensitive than the Rosenthal-Fode study to any *E* bias effects that were present. The implication is invalid. The overall parametric analysis that was performed in the Barber et al. investigation was based on 51 *Es* (testing 501 *Ss*) and was at least as sensitive to *E* bias effects as the parametric analysis performed in the Rosenthal-Fode study which was based on 10 *Es* (testing 206 *Ss*).

Sex of experimenter. Most of the *Es* in the Rosenthal-Fode study were males whereas most of the *Es* in the Barber et al. investigations were females. Rosenthal's critique stated that this difference is important because "evidence has been presented (Rosenthal, 1966) that female *Es*, especially when contacting male *Ss*, are less likely to show expectancy effects. Indeed, they have been shown sometimes to obtain data significantly opposite to that expected." A recent paper (Barber & Silver, 1968) has analyzed this contention pertaining to female *Es* and has shown that it is misleading. Although there is some evidence indicating that female *Es*, especially when testing male *Ss*, show reduced or negative expectancy effects (Rosenthal, Persinger, Mulry, Vikan-Kline, & Grothe, 1964), there is also contradictory evidence. For instance, a study by Silverman (1966) showed that, as compared to other combinations of *E*'s sex with *S*'s sex, female *Es* testing male *Ss* obtain results more in line with their expectancies (not less in line or opposite to

¹ Preparation of this paper was aided by Grant MH-11521 from the National Institute of Mental Health, United States Public Health Service.

² Requests for reprints should be sent to Theodore X. Barber, Medfield State Hospital, Medfield, Massachusetts 02042.

their expectancies, as Rosenthal's critique contends).

Let us look more closely at Rosenthal's contention that the Barber et al. investigations included too many female *Es* and that this might have made a difference. Did it actually make a difference? That is, when the data of the Barber et al. studies are analyzed to determine the effects of *Es'* sex, do we find that female and male *Es* obtained different ratings from their *Ss*? Not at all! Under each of the expectancy conditions, female and male *Es* obtained practically identical ratings (all *Fs* pertaining to the effects of *Es'* sex were less than 1.0). Did female *Es* testing male *Ss* bias their *Ss* less than other *E-S* combinations? Not at all! The relevant interaction (*Es'* Sex \times *Ss'* Sex \times Treatments) yields an *F* much smaller than 1.0, which can be expected to occur by chance around 50 times in 100.

Experimenter status. In the Rosenthal-Fode study 70% of the *Es* were undergraduates whereas in the Barber et al. investigations all of the *Es* were undergraduates. Also, the *Es* in the Rosenthal-Fode study were enrolled in an undergraduate experimental psychology laboratory whereas the *Es* in the Barber et al. studies were enrolled in an introductory psychology course. Rosenthal's critique pointed out, in the same way as the paper by Barber et al. had pointed out, that these differences give rise to a discrepancy in the relative status of the *Es*; that is, the *Es* in the Rosenthal-Fode study tended to be of somewhat higher status than those in the Barber et al. studies. Rosenthal's critique also stated that higher status *Es* are more likely to be successful in biasing their *Ss'* responses. The latter statement, however, is not clearly in line with Rosenthal's own data (Rosenthal, 1966, pp. 241-245). Two studies have been conducted in Rosenthal's laboratories to determine whether high-status or low-status *Es* exert more bias. As Rosenthal (1966, pp. 244-245) himself clearly pointed out in his recent text, one of these studies (Laszlo & Rosenthal, 1967) showed unequivocally that low- rather than high-status *Es* exert more bias. The other study, a master's thesis supervised by Rosenthal (Vikan-Kline, 1962), appeared to indicate that high-status *Es* are "more successful at influencing their subjects to yield the desired data, but only among subjects contacted later in the experiment [Rosenthal, 1966, p. 242]." However, as Rosenthal (1966, pp. 242-245) clearly pointed out in his recent text, in the latter study the *Es'* status was confounded with several other variables—e.g., age and self-assurance—and the results could not be clearly attributed to differences in status. It appears that further research is needed

to pinpoint the effects of *Es'* status on the Experimenter Bias Effect.

Fortunately, the required "further research" that was mentioned in the preceding sentence has been conducted. Very recently, Wessler and Strauss (1968) presented a study which attempted to replicate the standard Rosenthal procedures and which was very similar to the Barber et al. studies. However, the Wessler and Strauss study differed from the studies by Barber et al. in that all *Es* were college seniors and their *Ss* were freshmen or sophomores. Although the *Es* in the Wessler and Strauss study were higher in status than their *Ss*, the results were practically identical with those obtained by Barber et al. Wessler and Strauss reported that *Es* expecting high (+5.0) ratings obtained mean ratings of -0.52 whereas those expecting low (-5.0) ratings obtained mean ratings of -0.64. These differences are very far from significant ($p > .50$).

In brief, it appears that (a) the statement found in Rosenthal's critique, that high- rather than low-status *Es* exert more bias, is unequivocally contradicted by one study from Rosenthal's laboratory and is not clearly supported by another, and (b) when an attempt is made to replicate the basic Rosenthal procedures using *Es* who are relatively high in status, results are obtained which are essentially the same as those obtained by Barber et al.

POSTMORTEM DATA ANALYSES

The basic data that were analyzed in the Rosenthal-Fode study (and in almost all other studies carried out by Rosenthal and associates with the person-perception task) were the *mean ratings* obtained by the *Es*. Rosenthal and Fode performed a *parametric* test on these mean ratings. To adhere as closely as possible to the procedures used in the Rosenthal-Fode study, it was incumbent upon Barber et al. to perform a *parametric* analysis on the *mean ratings* obtained by their *Es*. This appropriate statistical analysis, which yielded an overall *F* less than 1.0, showed no evidence that *Es* obtain results in line with their expectancies. (Supplementary nonparametric analyses, which included a Kruskal-Wallis test and a chi-square test, were also performed on the *mean ratings* obtained by the *Es*. These analyses also failed to show an Experimenter Bias Effect.)

Instead of accepting the nonsignificant results that are obtained when the data of Barber et al. are analyzed in the same way that Rosenthal and Fode analyzed their data, Rosenthal's critique proposes that the Barber et al. experiments should be analyzed by a quite different method, namely,

TABLE 1

RATINGS OF PHOTOGRAPHS BY 501 Ss WITH COMPONENT CHI-SQUARES AND p VALUES

Photo ratings	Expectancy of E				
	+5.0	None	-5.0	χ^2	p
+3.00 and above	9	5	6	1.30	<i>ns</i>
+2.50 to +2.99	9	6	0	8.40	.02
+2.00 to +2.49	8	2	5	3.60	<i>ns</i>
+1.50 to +1.99	11	19	10	3.65	<i>ns</i>
+1.00 to +1.49	14	15	25	4.11	<i>ns</i>
+ .50 to + .99	14	17	28	5.51	<i>ns</i>
.00 to + .49	30	20	29	2.30	<i>ns</i>
- .50 to - .01	21	26	21	.73	<i>ns</i>
-1.00 to - .51	18	22	14	1.78	<i>ns</i>
-1.50 to -1.01	16	11	12	1.07	<i>ns</i>
-2.00 to -1.51	4	10	8	2.54	<i>ns</i>
-2.50 to -2.01	5	9	0	8.71	.02
-2.51 and below	8	5	9	1.18	<i>ns</i>
Total	167	167	167	44.9	.007

by a *nonparametric* chi-square test which is applied to the individual *Ss' ratings* and which does not take the individual *Es* into consideration. Although there are several objections we could raise here, let us accept Rosenthal's proposal to perform a chi-square test on the ratings of the 501 *Ss*. What is the logical way to partition these ratings for a chi-square test? The most logical cut is at the median. When the ratings are thus logically dichotomized above and below the median, we find that (a) more *higher* ratings are obtained when *low* ratings are expected (opposite to the results predicted by the *E* bias hypothesis) but (b) these differences are not significant ($\chi^2 = 4.36$, $df = 2$, $p > .10$).

Since the ratings given by the 501 *Ss* can be partitioned in many other ways, many additional chi-square tests can be performed. Of course, if we continue to carry out postmortem statistical analyses, we may expect by chance to find some "significant" results (Hays, 1963, Chap. 14). Table 1 in Rosenthal's critique presented one of the many additional chi-square tests that can be performed on the Barber et al. data. Rosenthal's overall chi-square test, which was performed on the *Ss' ratings* partitioned into 13 discrete intervals, yielded a significant p , indicating that one or more of the 13 intervals deviated from chance. However, Rosenthal's critique failed to make clear exactly where the deviations from chance expectations were localized in the array of data. To determine where the ratings deviated from chance, it is necessary to look at the component chi-squares that were computed for each of the 13 rows in Rosenthal's Table 1. These chi-squares and their associated p values, which Rosenthal's critique failed to present, are listed in Table 1 in this paper. The table shows the following:

1. *Contrary to the E bias hypothesis*, the *Es' expectancies* did not significantly affect 11 of the 13 sets of ratings (ratings of -2.51 and below, -2.00 to -1.51, -1.50 to -1.01, -1.00 to -.51, -.50 to -.01, .00 to 0.49, +.50 to +.99, +1.00 to +1.49, +1.50 to +1.99, +2.00 to +2.49, and +3.00 and above).

2. *Contrary to the E bias hypothesis*, a significantly smaller number of *Ss* tested by *Es* expecting low (-5.0) ratings gave relatively low ratings between -2.50 and -2.01.

3. In comparison with *Ss* tested by *Es* expecting low (-5.0) ratings, a significantly larger number of *Ss* tested by *Es* with no expectancies and by *Es* with expectancies for high (+5.0) ratings gave relatively high ratings between +2.50 and +2.99.

It needs to be emphasized here that (a) the appropriate analysis of variance applied to the

mean scores obtained by the 51 *Es* in the Barber et al. studies yielded nonsignificant results ($F < 1.0$) and (b) various supplementary nonparametric tests (Kruskal-Wallis tests and chi-square tests) also yielded nonsignificant results. Consequently, the three sets of "findings" presented immediately above, which are due to an additional postmortem chi-square test performed on an arbitrary partitioning of the *Ss' ratings* into 13 intervals, can be viewed, at best, as three tentative propositions that require experimental validation. Since these three tentative propositions are very specific and have very little if any general interest, it is doubtful that investigators will bother to test them. However, if anyone wishes to validate these tentative propositions, he might note the following: (a) The first two contradict the *E* bias hypothesis which states that *Es* obtain results in line with their expectancies. (b) The third tentative proposition is partly inconsistent with the *E* bias hypothesis (in that *Es* expecting high ratings and those with no expectancies did not differ significantly in influencing the relatively high ratings between +2.50 and +2.99).

REFERENCES

- BARBER, T. X., CALVERLEY, D. S., FORCIONE, A., McPEAKE, J. D., CHAVES, J. F., & BOWEN, B. Five attempts to replicate the Experimenter Bias Effect. *Journal of Consulting and Clinical Psychology*, 1969, 33, 1-6.
- BARBER, T. X., & SILVER, M. J. Pitfalls in data analysis and interpretation: A reply to Rosenthal. *Psychological Bulletin*, 1968, 70(6, Pt. 2) 48-62.

- HAYS, W. L. *Statistics for psychologists*. New York: Holt, Rinehart & Winston, 1963.
- LASZLO, J. P., & ROSENTHAL, R. Subject dogmatism, experimenter status, and experimenter expectancy effects. Department of Social Relations, Harvard University, 1967. (Mimeo)
- ROSENTHAL, R. *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts, 1966.
- ROSENTHAL, R. On not so replicated experiments and not so null results. *Journal of Consulting and Clinical Psychology*, 1969, 33, 7-10.
- ROSENTHAL, R., & FODE, K. L. Three experiments in experimenter bias. *Psychological Reports*, 1963, 12, 491-511.
- ROSENTHAL, R., PERSINGER, G. W., MULRY, R. C., VIKAN-KLINE, L., & GROTHE, M. Changes in experimental hypotheses as determinants of experimental results. *Journal of Projective Techniques*, 1964, 28, 465-469.
- SILVERMAN, I. The effects of experimenter outcome expectancy on latency of word association. Paper presented at the meeting of the Eastern Psychological Association, New York, April 1966.
- VIKAN-KLINE, L. The effect of an experimenter's perceived status on the mediation of experimenter bias. Unpublished master's thesis, University of North Dakota, 1962.
- WESSLER, R. L., & STRAUSS, M. E. Experimenter expectancy: A failure to replicate. *Psychological Reports*, 1968, 22, 687-688.

(Received July 22, 1968)

REFLECTIONS ON REPLICATIONS AND THE EXPERIMENTER BIAS EFFECT

LEON H. LEVY¹

Indiana University

The value of replications in psychology, the need to distinguish between the probability of data and of hypotheses, and the implications of alternative conceptualizations of the Experimenter Bias Effect (*EBE*), were each discussed in the light of the failure reported by Barber, Calverly, Forgione, McPeake, Chaves, and Bowen (1969) to replicate earlier findings by Rosenthal and Fode (1963). The position was taken that the value of further research on the *EBE* will depend upon the prior development of a situational taxonomy and a theoretical formulation in which the effect may be conceptualized in terms of a broader class of phenomena.

It is axiomatic among novelists, particularly of the existentialist persuasion, that character is revealed in conflict and crises, and this seems no less true when we turn to questions of the nature of scientific inquiry and thought (Kuhn, 1964). Thus, it would be unfortunate were we to focus only on questions of whether the study by Barber, Calverley, Forgione, McPeake, Chaves, and Bowen (1969) should be counted as a serious and legitimate failure to replicate the results of the Rosenthal and Fode (1963) experiment demonstrating an Experimenter Bias Effect (*EBE*), or whether this effect might be revealed by a more fine-grained, post hoc analysis of their data, and failed to take advantage of the opportunity afforded by their study to reflect upon some of the broader and more fundamental issues which are raised by it.

These issues are joined once the question is asked: What are the implications of the failure of Barber et al. to demonstrate an *EBE* in their several replications? The first concerns the value of replications; the second concerns the relationship between data and hypotheses; and the third concerns the implications of the alternative stances which might be taken toward the *EBE*. These are certainly not independent of each other, nor are they specific to the *EBE*, but they are especially relevant at this time since it is apparent that the *EBE* has become an increasingly popular topic and bids fair to rival anxiety as a vehicle for the production of theses and journal articles. And in psychology, as elsewhere, there seems to be an inverse relationship between the popularity of activity and the amount of thought given to its meaning or value.

The Value of Replications

The perfect replication is a fiction, and I shall take the heretical position that this is just as well. Therefore, let us stipulate that the Barber et al. experiment differs from Rosenthal and Fode's in several details. Some of these are listed by Barber et al. But does this necessarily disqualify their findings as evidence vis-à-vis those of Rosenthal and Fode concerning the *EBE*? For obvious reasons, no experiment can ever duplicate another in every detail, and so this question turns on whether the variations between them were trivial or important. Thus, we have the choice of either of two conclusions: (a) Barber et al. failed to establish one or more of the conditions necessary for the demonstration of the *EBE*; or (b) their procedural variations were trivial and the *EBE* is unreliable and/or not very pervasive.

But this choice is an impossible one, for it requires either a theory which states the parameters involved in the *EBE* or a body of systematic research from which these parameters might be induced. While it seems clear that the former condition does not exist in the case of the *EBE*, for reasons which we cannot go into in the space available, it might also be argued that the extensive research which has been performed in the name of the *EBE* cannot be considered systematic. Thus, whether the findings of Barber et al. can be taken as contradictory to those of Rosenthal and Fode is a moot question, and this, I would suggest, will be found true wherever replications are attempted of experiments dealing with phenomena which are not embedded either within some theoretical framework or extensive body of systematic research. Conversely, it might be argued that any dispute over whether an experiment qualifies as a replication of another may

¹Requests for reprints should be sent to Leon H. Levy, Department of Psychology, Indiana University, Bloomington, Indiana.

be taken as *prima facie* evidence of the weakness of the conceptual and empirical underpinnings of the phenomenon in question. And so long as this is the case, it would seem that unless we are interested only in the existence of a phenomenon in a particular situation (which is rarely, if ever, the case in science, and certainly not so in the case of the *EBE*) we must ask whether replications are worth the effort. Can we not employ our resources more profitably?

These questions also merit consideration in instances where phenomena are better endowed conceptually and empirically than the *EBE*. For, unless there are serious reasons to doubt the validity of the data reported by an investigator, it would seem wiser to invest our time in exploring the generalizability of the theory concerning the phenomenon through the systematic variation of its parameters and experimental procedures in a manner analogous to that represented by the multitrait-multimethod matrix proposed by Campbell and Fiske (1959) for the study of construct validity. Such experiments would permit an evaluation of the reliability and pervasiveness of the phenomenon in question while at the same time extending our theoretical understanding of it. Since variation from one experiment is inevitable, why not make the most of it? Surely when practice departs from precept as radically as it does in the matter of replications, so that an investigation such as that of Barber et al. is the exception rather than the rule, it may be time to examine the wisdom of the precept rather than lament its neglect. And as a start I would suggest that we consider how much more information was gained by Barber et al. through their second replication, or their third, fourth, or fifth. Measured against information gained, it would seem that these five replications were at least four too many; concerning the relationship between replications and the acquisition of knowledge, they suggest the aptness of the aesthetic dictum: *less is more*.

The Relationship between Data and Hypotheses

The demonstration of the *EBE* in the Rosenthal and Fode experiment and others might be likened to a wife's discovery of a strange shade of lipstick on her husband's shirt collar. Whether or not she had suspected him of infidelity in the past, it seems unlikely that any number of unblemished shirt collars in the future will completely remove her concerns about his conjugal constancy. Thus it seems unlikely that the five failures to reject the null hypothesis by Barber et al. will in itself convince many of the rarity of the *EBE* or that it need not be taken into account in the design and interpretation of experiments. Nor should it. Why this should be so

becomes apparent when we recognize the distinction, which is often ignored, between the probability of data and of hypotheses, and the difficulties entailed in drawing inferences from significance tests and null hypotheses. Since there are several excellent discussions of these issues in the recent literature (Bakan, 1966; Edwards, Lindman, & Savage, 1963; Rozeboom, 1960), I shall do no more than point out that they are made most timely by the Barber et al. paper.

The Implications of Alternative Views of the EBE

The *EBE* might be regarded either as a methodological problem or as a member of a class of phenomena of intrinsic interest to psychology. In the former case, it takes its place along with response biases and demand characteristics as, what we might call, a "spoiler variable," in that it spoils the image of the rigor and purity of psychology's data collection procedures and complicates the design and interpretation of research by requiring that an "irrelevant" dimension—the *E*—be taken into consideration. Regarded as an exemplar of a class of phenomena such as interpersonal influence or the effects of expectancies and values on perception and behavior, the *EBE* might be considered a convenient vehicle for the investigation of the processes involved in these phenomena and as a medium for testing hypotheses and theories about them. Each of these stances, however, entails a different set of considerations and different order of priorities so far as research and development are concerned.

Regarding the *EBE* as a methodological problem leads to considerations of its pervasiveness, how much of the variance in experiments might be attributed to it, and how it might be controlled or eliminated. Although Rosenthal (1966) has suggested a number of means by which this last concern might be met, the decision to incur the additional costs involved in adopting them cannot be made independently of the first two considerations. However, the information which would be relevant to these prior considerations is dependent, in turn, upon the existence of a taxonomy of situations (tasks or experiments). For without such a taxonomy and the sampling of situations it would permit, it is impossible to make any generalizations about the seriousness of the *EBE* as an epistemological threat to psychology on the basis of successes or failures in demonstrating it in particular instances. But although the value of such a taxonomy has long been recognized, and extends far beyond the problem of *EBEs* (Rotter, 1960), one does not presently exist. Thus we must ask whether it would not be wiser to give precedence to the long overdue development of this taxonomy rather than to addi-

tional demonstrations of the *EBE*, however startling and impressive these might be. Certainly, as matters stand at present, it should be obvious that there is no way of assessing the validity of any assertions about the generality or seriousness of the *EBE* as a methodological problem.

While the perverse gratifications which might derive from demonstrations of the *EBE* in its role as spoiler variable are not to be minimized, and regardless of its importance in this role, by far the more rewarding scientific stance toward the *EBE* is likely to be as an exemplar of some theoretically important class of psychological phenomena. The low epistemic yield of the five experiments reported by Barber et al. clearly reveals the wastefulness of research conducted on phenomena which are devoid of any theoretical conceptualization. For had the *EBE* been placed in some theoretical context, these five failures to demonstrate it might have been turned to good account in evaluating one or more theoretical propositions concerning its nature and the processes involved in it, or in deciding on directions for future research. It should be noted also, that if the *EBE* is conceived of as an exemplar of some class of phenomena, it becomes irrelevant whether it is pervasive or rare; the *EBE* could be exotic and still of importance if the class of phenomena were important and if it was of value in furthering our understanding of these phenomena.

These *ifs* point to the kinds of questions which must be asked if the *EBE* is to be taken seriously from a theoretical standpoint. They require, again, a better conceptualization of the situations in which the *EBE* is manifested than presently exists. And, paradoxically, they also require a high degree of development in the theory and knowledge concerning the phenomena allegedly exemplified by the *EBE* and which the *EBE* is intended to further illuminate. For, as Bordin (1965) has pointed out in a different context, it is only as a phenomenon is more fully understood that it can safely be studied in the laboratory by means of simplifications and analogues without the fear that these may be irrelevant to it. This means, at a minimum, that if the *EBE* is to be regarded as a vehicle for the study of theoretically important processes and phenomena, such as unintended covert communication, as Rosenthal (1967a) has recently proposed, it cannot be the sole means of doing so. It also means that hypotheses based upon *EBE*-type experiments must be tested for their generality in other experiments involving the phenomena in question, and vice versa.

Rosenthal (1967b) has observed that as re-

sponse variance diminishes so does variance in the *EBE* (p. 621). If this is the case, since two of the most effective ways of reducing (unwanted) response variance are through better experimental controls and more valid and comprehensive theories, it is appropriate to ask whether these should not be our first order of business, regardless of the stance taken toward the *EBE*. For if our reasoning is correct, this shift of emphasis away from the *EBE* to more fundamental methodological and substantive concerns should have the interesting consequences of reducing the *EBE* as a methodological threat while at the same time helping us to understand and use it more wisely in the furtherance of psychological knowledge.

REFERENCES

- BAKAN, D. The test of significance in psychological research. *Psychological Bulletin*, 1966, 66, 423-437.
- BARBER, T. X., CALVERLY, D. S., FORGIONE, A., McPEAKE, J. D., CHAVES, J. F., & BOWEN, B. Five attempts to replicate the experimenter bias effect. *Journal of Consulting and Clinical Psychology*, 1969, 33, 1-10.
- BORDIN, E. S. Simplification as a strategy for research in psychotherapy. *Journal of Consulting Psychology*, 1965, 29, 493-503.
- CAMPBELL, D. T., & FISKE, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, 56, 81-105.
- EDWARDS, W., LINDMAN, H., & SAVAGE, L. J. Bayesian statistical inference for psychological research. *Psychological Review*, 1963, 70, 193-242.
- KUHN, T. S. *The structure of scientific revolutions*. Chicago: University of Chicago Press, 1964.
- ROSENTHAL, R. *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts, 1966.
- ROSENTHAL, R. Covert communication in the psychological experiment. *Psychological Bulletin*, 1967, 67, 356-367. (a)
- ROSENTHAL, R. Psychology of the scientist XXIII. Experimenter expectancy, experimenter experience, and Pascal's wager. *Psychological Reports*, 1967, 20, 619-622. (b)
- ROSENTHAL, R., & FODE, K. L. Psychology of the scientist: V. Three experiments in experimenter bias. *Psychological Reports*, 1963, 12, 491-511.
- ROTTER, J. B. Some implications of a social learning theory for the prediction of goal directed behavior from testing procedures. *Psychological Review*, 1960, 67, 301-316.
- ROZEBOOM, W. W. The fallacy of the null-hypothesis significance test. *Psychological Bulletin*, 1960, 57, 416-428.

(Received June 13, 1968)

WHAT PATHOLOGICAL DIMENSIONS LIE BEYOND THE NORMAL DIMENSIONS OF THE 16 PF? A COMPARISON OF MMPI AND 16 PF FACTOR DOMAINS¹

RAYMOND B. CATTELL

AND

LAURA SPECHT BOLTON²

University of Illinois

Stanford University

A set of mutually exclusive, exhaustive parcels of MMPI items and the 16 personality factor scales from the Cattell 16 PF were factored together on 217 normal and 40 abnormal adult Air Force men. The analysis of the resulting 64 variables (45 MMPI parcels and 19 16 PF scales) yielded 21 factors, of which 13 are clearly identifiable as factors in the 16 PF and 4 are clearly MMPI factors. MMPI data were scored also for the traditional 14 clinical scales, and their predictability from 16 PF source trials examined in terms of the squared multiple correlations. The 7-factor space typically represented by the MMPI clinical scales was found to lie within the dimensionality of the Cattell 16 PF. Furthermore, the total prediction of the set of MMPI clinical scales from a single form (A) of the 16 PF, uncorrected for attenuation, is represented by the canonical correlation of .8.

The 16 PF claims, by the operational working out of the personality sphere concept (Cattell, 1946, 1957) in the basic research underlying it, to represent the main dimensions of the normal personality. There are, it is true, some seven further factors (Cattell, 1957) of small variance and rather unstable reappearance which research needs to pursue, but with one exception these also appear "normal."

What the clinician recognizes as abnormality can be either a statistically deviant score on a normal factor—or, alternatively, a score on some factor the very existence of which is demonstrable only in pathological manifestations, as is recognized on some MMPI scales.

To ask how much of the MMPI response material is pathological in this second sense requires that we first ask how much of the variance in such behavior can be accounted for by deviation on already known normal

factors. A joint factoring of the MMPI and 16 PF has therefore long been overdue, as many practicing psychologists have pointed out. The research here reported has as its immediate objective such a comparison of the factor space of the MMPI and the 16 PF. To appreciate all features of its design, however, one must recognize that it is also part of a series of researches (Cattell & Bjersted, 1967)³ to provide the theoretical basis for the Clinical Analysis Questionnaire to be added to the 16 PF. An instrument thus adding entirely pathological to normal factors would offer the clinician a means of handling diagnosis as well as the general personality of the individual.

In designing the experiment for the comparison, one must recognize that the MMPI was designed to differentiate among persons exhibiting various clinically defined surface traits (i.e., syndromes, formally and psychometrically correlation clusters) while the 16 PF was designed to measure source traits (replicable simple structure factors). Actually, there appears still to be no published basic research to show that the MMPI scales are constituted by discrete correlation clusters and the conclusion of correlational investiga-

¹ This research was supported in part by Public Health Research Grant MH 1733-09 from the National Institute of Mental Health and by the National Science Foundation through an undergraduate fellowship in support of the junior author. The writers wish to express their indebtedness to John Nesselrode and Jack Ford for extensive computer assistance, to J. Wiggins and L. Triandis for valuable comment, and to the chief psychiatrist and the administration of Chanute Air Force Base for providing Ss.

² Now at 155 Rose Street, Metuchen, New Jersey 08840.

³ R. B. Cattell and K. Delhees. Personality factors in the pathological domain, and the construction of the 16 PF Clinical Analysis Questionnaire, in preparation.

tions in this area (e.g., Wiggins, 1962), has indicated that putting together items which correlate adequately with possession of a designated clinical syndrome does not itself guarantee adequate homogeneity in such a scale. For this reason, and because of obvious technical factor-analytic blunders, discussed below, no published factor analysis of the MMPI has yet supplied the answer to the question we now raise about its primary factor and cluster structure.

Welsh (1952) performed a centroid analysis of shortened nonoverlapping versions of the MMPI clinical scales. However, this study sampled less than one-third of the total MMPI item space. Comrey (1957a, 1957b, 1957c; 1958a, 1958b, 1958c, 1958d, 1958e, 1958f, 1958g) performed extensive factor analyses of individual items within given MMPI conventional scales, with cross-cultural validation (Comrey & Nencini, 1961), rotating to a varimax solution. However, none of these investigations covered more than 12% of the total MMPI space. Other investigations, such as those of Wiggins (1966) and Mees, Gocka, and Holloway (1964) have examined MMPI factor structure with respect to certain item characteristics, such as content and social desirability.

With the above-mentioned exceptions, investigators have factored the scores on conventional or overlapping scales, typically obtaining some agreement on only five or fewer factors. Practically none have met the ordinary technical demands for tests of the number of factors, proof of rotation to maximal simple structure, objective testing of alleged matches with other studies, and especially the avoidance of the fatal weakness, discussed in detail by Shure and Rogers (1965), of having algebraic dependencies, through item overlap, in the scales factored. (Among others, see Finney, 1966; Kassebaum, Couch, & Slater, 1959; Lingo, 1960; La Forge, 1962; Tyler, 1951; and Wheeler, Little, & Lehner, 1951).

The 566 items collected by the constructors of the MMPI probably represent as comprehensive a sampling of pathological behavior as the 564 in the 16 PF do of normal behavior. At any rate, we would accept them as the best sampling of the domain yet available. But to reveal the real dimensionality of

the MMPI domain it is necessary to forget the present cluster scales, possibly arbitrary, and factor the full roster of items themselves. In effect this is what the present research, in conjunction with a coordinated study by Eber and Fowler,⁴ sets out to accomplish.

Incidentally, in undertaking this factor analysis of parcels without regard to the structure imposed by the MMPI test constructors in what we may, for brevity, call the "conventional cluster scales" (including the various additional scales added by other clinicians), it is not implied that measuring patients by syndromes (as technically defined by surface trait, cluster search operations Cattell, 1946, 1957; Cattell & Coulter, 1966) is unnecessary. On the contrary, in psychiatric literature, Cattell (1966a, 1967) has advocated the special value of depth psychometry, that is, a planned simultaneous measurement of surface traits, for example, by the MMPI, and of source traits, for example, by the 16 PF, MAT, etc. If we are dealing with a surface-trait and a source-trait instrument, both occupying the same factor space, then the depth psychometry technique would give the clinician the advantage of understanding two patients with identical scores on a given syndrome scale in the light of the different source-trait structure underlying each. This is a first step from descriptive to causal analysis.

To explore and illustrate the possibilities of depth psychometry while investigating the primary factor structure, two distinct experiments and factorings of MMPI-16 PF data have been undertaken here. The first takes the MMPI data as conventionally scale-scored and relates it to 16 PF source traits while the second partitions the MMPI, regardless of conventional scoring, to find what is contained in a set of mutually exclusive, exhaustive parcels from it.

EXPERIMENT I: MMPI SURFACE-TRAIT AND 16 PF SOURCE-TRAIT RELATIONS

For the proper determination of the factor structure, it was considered desirable, in administering the two scales together, not to use

⁴H. W. Eber, Alton State Hospital, Alton, Illinois, personal communication, April 1966.

the usual student groups, but to find a more representative section of the population, and one, moreover, containing a reasonable percentage of "clinical" cases. Accordingly, the MMPI and the 16 PF, the latter as Form A, were administered to 259 Air Force men, including, for "pathological" variance, about 40 cases listed by the psychiatrists as being in some need of psychiatric help.

The correlations between the two sets of scales—MMPI and 16 PF—are shown corrected for attenuation in Table 1. In order to obtain the closest statement of the real relations of the scales' concepts per se, it is necessary, of course (since we took only a very brief 10-item score of most 16 PF scales, in using Form A only), to correct for attenuation. The uncorrected correlations are available from the National Auxiliary Publications Service^a (see Table A) which also presents the correlation matrices among (a) the MMPI scales alone and (b) the 16 PF scales alone.

A factor analysis of MMPI intercorrelations alone, which we shall not present here because we are not interested further in a priori MMPI scale scores as such, shows by

^a For supplementary material, order NAPS Document NAPS-00121 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001; remitting \$1.00 for microfiche or \$3.00 for photocopies.

the scree test (Cattell, 1966c) that they contain either seven or eight factors. This finding is compatible with earlier studies (e.g., La Forge, 1962; Messick & Jackson, 1961; and others) which centered on five or six but typically did not iterate communalities, apply recent tests (or any true tests) for the number of factors, or use scales cleared of algebraic dependence. However, even if we accept, as a compromise, that the number is seven, we can see why it is impossible for some of the 16 PF factors, for example, A, B (intelligence), L, M, Q₁, and Q₂ to have more than trivial or insignificant correlations (Table 1) with these surface traits. The simplest ways to give expression to the amount of overlap of the two systems are (a) to ask how much of the variance of the scales of one can be predicted from the other, that is, by obtaining a series of squared multiple Rs in both directions, or (b) to calculate the canonical correlation.

For most purposes a practitioner will be more interested in the first of these, and in the direction of asking how much of the MMPI surface-trait scores he can predict from the 16 PF scale scores. These multiple correlations were corrected for attenuation to ameliorate the unreliabilities of both sets of scales, as reported in the respective handbooks of the 16 PF and the MMPI. Table 2 is obtained by taking the corrected correlations of Table 1, converting them into weights (for estimating

TABLE 1

CORRELATIONS BETWEEN MMPI AND 16 PF SCALES (CORRECTED FOR ATTENUATION ON BOTH)

Scale	A	B	C	E	F	G	H	I	L	M	N	O	Q ₁	Q ₂	Q ₃	Q ₄
<i>A</i>	-.13	-.26	-.40	.17	.10	-.31	-.26	.06	.20	.10	-.35	.63	-.18	.04	-.52	.62
<i>L</i>	-.07	.01	.29	-.25	-.02	.14	.01	-.20	-.21	.03	.19	-.22	-.06	-.03	.19	-.13
<i>F</i>	-.07	-.38	-.41	.21	-.15	-.68	-.10	.03	.23	.19	-.49	.32	.20	.20	-.62	.39
<i>K</i>	-.04	.11	.42	-.27	.02	.26	.09	-.23	-.36	-.13	.35	-.43	-.02	-.06	.46	-.36
<i>Hs</i>	.01	-.33	-.45	.13	-.04	-.46	-.09	-.01	.24	.29	-.39	.52	-.10	.12	-.48	.48
<i>D</i>	-.13	-.10	-.23	-.33	-.27	-.15	-.25	.00	.12	.16	-.07	.20	-.20	.21	-.18	.24
<i>Hy</i>	.01	.30	.20	-.29	-.04	.01	.03	-.05	-.13	.23	.11	-.14	-.07	-.07	.08	-.05
<i>Pd</i>	-.07	-.12	-.14	-.02	.01	-.37	-.05	-.08	.16	.27	-.04	.24	-.11	.07	-.30	.36
<i>Mf</i>	.05	.10	.08	-.33	-.14	.11	-.05	.14	-.26	.17	.05	-.10	.03	.03	.22	-.06
<i>Pa</i>	-.08	-.24	-.15	-.08	-.11	-.33	-.01	-.05	.08	.20	-.18	.12	.01	-.04	-.33	.21
<i>Pt</i>	-.12	-.37	-.64	.02	-.15	-.47	-.29	.10	.45	.25	-.29	.68	-.20	.09	-.74	.71
<i>Sc</i>	-.01	-.47	-.53	.16	-.09	-.59	-.13	.12	.35	.23	-.51	.54	.02	.10	-.72	.56
<i>Ma</i>	-.01	-.15	.19	.15	.20	-.25	.13	-.11	.14	.11	.14	-.09	.08	-.12	-.14	.12
<i>Si</i>	-.33	-.10	-.17	-.49	-.42	-.09	-.44	-.13	.04	.07	-.04	.26	-.10	.43	-.14	.24

TABLE 2

PREDICTIONS (MULTIPLE CORRELATIONS) OF MMPI SURFACE TRAIT SCALES FROM 16 PF SOURCE TRAIT SCALES (WITH VARIOUS ATTENUATION CORRECTIONS)

Scale	Corrected for MMPI scale unreliabilities only	Corrected for 16 PF scale unreliabilities only	Corrected to ideal scales for both	Uncorrected
Anxiety	49	64	68	58
Lie	27	35	40	32
Validity	46	58	73	51
K Correction	37	49	55	43
Hypochondriasis	42	55	62	49
Depression	32	36	47	32
Hysteria	28	29	43	26
Psychopathic	32	32	48	29
Masculinity-Femininity	25	33	38	29
Paranoia	30	36	48	32
Psychasthenia	55	70	78	63
Schizophrenia	48	66	75	58
Hypomania	33	35	48	32
Social Introversion	47	52	64	47

MMPI scores from 16 PF scales) and calculating multiple R s in the usual way. The weights are obtained by the matrix equation

$$W = C_r R^{-1}$$

where W is the vector of 16 weights to be attached to the 16 PF factor scale scores in estimating the given MMPI factor, C_r is the correlation matrix among the 16 PF scales. The latter turned out, incidentally, to be very much the same in this sample as in other samples (Cattell & Eber, 1967) and when factored⁶ gives the typical second-order structure.

It becomes evident that some MMPI scale scores, notably schizophrenia, anxiety, psychasthenia, and social introversion, can be predicted with appreciable efficiency from the 16 PF. A recent result by Eber (1966) independently supports these conclusions in that he obtains the appreciable multiple R of .7 in predicting from the 16 PF the psychotic-nonpsychotic decision (mainly on a schizophrenic group) as given by the usual judgments on MMPI data. The results are also consistent with the findings that the directions in which the 16 PF has been found most potent in clinical practice (Cattell & Scheier,

1961; Cattell, Tatro, & Komlos, 1964) are those of neuroticism (including anxiety and psychasthenia) and separation of the schizophrenias.

The converse prediction—MMPI to 16 PF scales—may be left to the reader,⁷ but the total overlap may be appreciated also from the first components of the canonical correlation which we have worked out from NAPS Table A, showing the possible prediction of MMPI from 16 PF space to be .80. However, these figures and the multiple R s also say that an appreciable amount in the MMPI is unaccounted for by the 16 PF (and vice versa) and this fits our later evidence that pathological domains in the latter still remain to be accounted for by the new 16 PF Clinical Analysis Questionnaire.

EXPERIMENT II: THE INHERENT COMMON DIMENSIONALITY OF THE MMPI AND 16 PF

The second and major experiment in our investigation aimed to evaluate the common dimensionality in the inherent item material, setting aside that scale development in the MMPI just considered, which, to a psychometrist or personality theorist, represents an odd history of piecemeal approaches and expediency. On contemplating the task of factoring 566 MMPI items one can, however, understand the previous reluctance to come to grips with the problem in more general theoretical terms.

Two possible approaches presented themselves to us; one—building special, large computer programs for rotation—was proposed by Eber and Fowler, while the present writers resorted to the parceled factor-analysis method (Cattell, 1960) which has proved effective before with large problems. By this latter, the totality of the MMPI items were grouped into 45 small parcels, averaging a dozen items in length. Complete contents of these parcels in terms of MMPI item numbers are listed in Table 3. We repeat for emphasis that these parcels were derived from scales with cor-

⁷ It can be seen, however, that the dimensionality of the pathological surface traits is only about half that of the 16 PF, so that the MMPI scales were not able to predict, for example, A (schizothymia), B (intelligence), Q₁ (radicalism-conservatism), and so on, in the normal personality sphere.

⁶ R. B. Cattell and J. Ford. An integration, by congruence coefficients, of the principal researches on second order personality factor structure in Q-data, in preparation.

TABLE 3
ITEM CONTENT OF MMPI PARCELS

Variable	Descriptive title of parcel	MMPI item composition of parcel in terms of item numbers in the MMPI test booklet (1951)
20	Welsh's Anxiety scale (All true)	41, 67, 94, 305, 356, 382, 396, 414, 418, 443, 465, 499, 511, 518, 544, 584.
21	Depression (true)	5, 13, 52, 86, 104, 130, 158, 159, 182, 193, 236, 259.
22	Depression (false)	46, 58, 64, 88, 95, 98, 131, 145, 154, 191, 207, 233, 242, 270, 272.
23	Faking to look bad (true)	14, 31, 34, 35, 40, 48, 49, 50, 53, 56, 66, 85, 121, 139, 146, 184, 197, 200, 202, 205, 206, 209, 211, 218, 227, 256, 291.
24	Faking to look bad (false)	20, 83, 113, 185, 199, 257.
25	Hypochondriasis (true)	23, 29, 43, 62, 72, 108, 114, 125, 161, 189, 273.
26	Hypochondriasis (false)	2, 9, 51, 55, 63, 68, 103, 153, 155, 175, 188, 190, 192, 230, 243, 274.
27	Hysteria (true)	10, 32, 44, 47, 76, 179, 186, 238, 246, 258.
28	Hysteria (false)	6, 12, 71, 128, 129, 136, 147, 162, 174, 213, 234, 265, 279, 292.
29	K: Correction for faking (all false)	138, 142, 160, 170, 171, 296, 374, 397, 398, 406, 461, 482, 502.
30	L: Faking to seem good (false)	15, 30, 45, 60, 75, 90, 150, 165, 195, 225, 295.
31	Hypomania (true)	11, 59, 73, 100, 109, 143, 167, 181, 222, 228, 232, 240, 250, 271, 277, 298.
32	Hypomania (false)	101, 105, 111, 119, 148, 166, 180, 267, 289.
33	Personal and emotional sensitivity (true)	217, 226, 278, 282, 299.
34	Personal and emotional sensitivity (false)	79, 99, 176, 198, 214, 254, 262, 264.
35	Female sexual identification or timidity (false)	19, 26, 28, 80, 89, 112, 120, 280.
36	Feminine occupational identification (true)	4, 25, 70, 77, 87, 92, 126, 132, 140, 149, 203, 204, 261, 295.
37	Denial of masculine occupations (false)	1, 81, 116, 144, 219, 221, 223, 283.
38	Paranoia (true)	16, 22, 24, 27, 123, 151, 275, 293, 338, 365.
39	Paranoia (false)	93, 117, 268, 313, 316, 319, 327, 347, 348, 294.
40	Psychopathic deviate (true)	42, 61, 84, 118, 215, 216, 224, 239, 244, 245.
41	Psychopathic deviate (false)	37, 82, 91, 134, 173, 183, 235, 237, 287, 96.
42	Psychasthenia (true)	304, 317, 336, 337, 340, 342, 343, 344, 346, 349, 351, 357, 358, 359, 360, 361.
43	Psychasthenia (false)	3, 8, 36, 122, 152, 164, 178, 329, 353.
44	Welsh's R (Repression) scale (false)	429, 445, 447, 468, 472, 516, 529, 550.
45	Schizophrenia (true)	168, 210, 297, 303, 312, 320, 324, 325, 334, 339, 345, 350, 354, 355, 363.
46	Schizophrenia (false)	17, 65, 177, 187, 196, 220, 276, 281, 306, 309, 322, 323, 330.
47	Social Introversion (true)	124, 172, 201, 377, 383, 411, 427, 436, 455, 473, 487, 549, 564.
48	Social Introversion (false)	33, 229, 391, 400, 415, 440, 446, 449, 450, 462, 469, 479, 505, 521, 547.
49	Edwards' Social Desirability (true)	7, 18, 54, 107, 163, 169, 371, 528.
50	Edwards' Social Desirability (false)	156, 241, 247, 252, 263, 269, 286, 301, 321, 335, 352, 424, 431, 439, 555.
51	Crown-Marlowe Social Desirability (true)	556.
52	Crown-Marlowe Social Desirability (false)	39, 135, 157, 255, 481.
53	Response set true (true)	310, 311, 326, 332, 341, 367, 370, 401, 405, 412, 421, 425, 430, 433, 435, 441, 452, 460, 463, 464, 467, 474, 478, 484, 485, 492, 494, 496, 486.
54	Response set false (false)	498, 504, 507, 508, 517, 519, 522, 523, 524, 525, 530, 532, 533, 534, 535, 537, 538, 539, 540, 541, 542, 543, 551, 557, 559, 560, 561, 563, 565.
55	Fundamental religion (true)	373, 413, 420, 476, 483, 488, 490.
56	Fundamental religion (false)	491.
57	Agree with generalized ext. standards (true)	302, 314, 364, 368, 378, 393, 409, 417, 438, 442, 448, 458, 466, 470, 493, 495, 510, 545, 548, 558.
58	Agree with generalized ext. standards (false)	437, 444, 459, 513, 514, 566.

TABLE 3—(Continued)

Variable	Descriptive title of parcel	MMPI item composition of parcel in terms of item numbers in the MMPI test booklet (1951)
59	Freedom from external control (true)	308, 315, 331, 333, 362, 372, 380, 381, 386, 387, 419, 422, 434, 456, 471, 506, 512, 536, 475.
60	Freedom from external control (false)	399, 423.
61	Mutuality (true)	375, 388, 390, 404, 408, 416, 477, 489, 500, 503, 509, 515, 527, 531, 554, 562.
62	Mutuality (false)	426, 432, 453, 454, 520, 552.
63	Expression of autonomy (true)	318, 366, 369, 376, 402, 407, 410, 428, 497, 501, 546.
64	Expression of autonomy (false)	328, 385, 392, 457, 526, 553.

responding labels, resulting from various previous investigations, in such a way as to preclude item overlap and maximize parcel homogeneity. Thus, as may be observed from Table 3, the parcel corresponding to Welsh's Anxiety Scale contains only 16 of the original items. The items were put together as far as possible with regard to item homogeneity, guided by content inspection and the way items were put together in the original clinical construction of the scales. Thus, some advantage was taken of clinical experience about homogeneity present in the scales, but the main dependence was on the work of Adams and Horn (1965) continued by Hundleby, to avoid any item overlap among the parcels.

True and false components of every MMPI parcel were separated to permit a check on the acquiescence response set. Thus, a variable labeled "true" or "false" in Table 3 consists of the subsets of items for which "true" or "false" responses, respectively, would yield a high score on the corresponding MMPI scale.

The plan now called for a factor analysis of the 45 parcels, from five to 20 items. From this, we would factor scores via parcels, and so obtain the correlations of all 566 items with the obtained factors. One should note that the success or lack of success in putting items in parcels with a better-than-chance degree of homogeneity is not crucial to success of the plan and does not fundamentally prevent the factor analysis having a correct outcome. For, in the last resort, even if the 566 items were thrown together in random groupings, the dimensionality of the variance among the parcels should be the same as the dimensionality among the items, except for trivial instances, as may be observed from the

principles of matrix algebra, upon which factor-analytic procedure is based. The adequacy of the "parceled" factor description of the domain is limited only by the fact that we cannot extract more factors than half of the number of variables in attempting to reach the full dimensionality. Since $n/2$ is still large—32—when all our variables are counted, there is slight danger that this could result in fewer dimensions being recognized than through factoring the items directly. The test of this is that when we ended by factoring (with the addition of 19 representatives of 16 PF scales to 45 MMPI parcels) a 64×64 matrix, the number of factors we obtained was almost exactly the same as that obtained by Eber and Fowler's coordinate study factoring the 566 items directly.

As Ss for this 64-variable battery we used virtually the same (2 missing) sample group of Air Force men as in the previous investigation, namely, 257 Air Force men (not pilot cadets), ages 19–25. For administration purposes, the Ss were divided into two roughly equal-sized groups, each completing the testing in two 2-hour sessions (morning and afternoon) on the same day. One group took the MMPI first and the 16 PF second, and the other group reversed this order. The general testing conditions prevailing were satisfactory, and since the Ss were guaranteed anonymity, we had the impression that suitable psychological conditions prevailed.

THE ANALYSIS

In proceeding to the correlation matrix, two possibilities arise: (a) to split each 16 PF scale into two parts, in order to guarantee emergence, by two markers, of each 16 PF

factor concerned at the primary level, or (b) to work with a somewhat more convenient smaller matrix and have better scale reliabilities, by leaving certain 16 PF scales in only as single scales (among the 45 MMPI parcels). The best solution seemed to leave the 16 PF as a single scale wherever the first investigation offered evidence that the 16 PF scale in question would certainly find some common variance with it in the MMPI material, but to put in two markers where this might be in doubt. Inspection of Table 1 showed some resonance in the MMPI could be expected for all scales except possibly N, the shrewdness dimension, Q₁, the radicalism-conservatism factor, and G, the super-ego dimension. Accordingly, the latter scales were split into two markers for each, but the remaining 13 factor scales were put in singly.

The Pearson product-moment 64×64 matrix was factored first by principal axis, inserting unities in the diagonals, in order to estimate the number of factors by the scree test (Cattell, 1966c). The results of the eigenvalue scree test in Figure A (see Footnote 5) indicate that 21 or 22 factors needed to be taken out. We accepted the former for economy and iterated 10 times upon which the communalities reached stability for 21 factors.

The unrotated V_o , the transformation matrix, and the correlation matrix of first-order factor procedure are deposited with the National Auxiliary Publications Service as Tables B, C, and D (see Footnote 5). The rotated matrix is the end result of several successive procedures: Maxplane (Eber, 1966), followed by a Procrustes clarification, followed by a long succession of Rotoplot rotations (see Figure B) to test that the highest possible simple structure had been reached. Since some who have been surprised by so much resolution of MMPI content into 16 PF space have suggested that the greater reliability of the 16 PF scales may have produced this result, it must be pointed out that, if anything, the MMPI had the better chance of statistically predominating. For (a) there were some 45 MMPI scales against only 19 16 PF subscales; (b) at the item level the predominance is even greater, 566 to 174; (c) the reliabilities of the two sets of scales (at least 20 parcels the same or greater item

length than 16 PF variables) are about the same, and (d) the 16 PF variables did not benefit from initial "far out" leverage, in the rotation, for, as the V_o matrix shows, the communalities of the MMPI scales are at least as high as those among the 16 PF scales; (e) the rotation is in any case finally fixed by the hyperplanes, not by high-loaded variables.

The hyperplane history appearing in Figure B (see Footnote 5) shows that the rotation followed the usual course from a moderate 45% of the variables in the ± 10 hyperplane to a number which is increasingly difficult to improve upon. The last four or five of the 27 rotations were persisted in without significant improvement (or loss) to make sure that the resolution had in fact reached a stable plateau. In fact, variation in any possible direction from our accepted position tended to produce a slight drop rather than an improvement. The hyperplane count for the final rotation rested at 79%, which is, by all standards in rotational experience in personality research, a highly adequate level.

The resultant reference vector matrix at simple structure is shown in Table E (see Footnote 5). Since the overall patterns lead to the conclusion that the majority of the factors are definable as 16 PF factors, the reference vectors have been arranged in the usual order of 16 PF factors for easy reference.

INTERPRETATION

By stating that the overall patterns revealed lead one to identify most as 16 PF factors, we do not mean that the 16 PF scale marker concerned necessarily has a really high loading, though it is (with two exceptions) statistically significant at $p < .01$ (Harris, 1964) and generally the highest in the column. These brief scales typically have a validity against the factor of about .6, but in the military samples of the kind here used, reliabilities run lower than usual and the validities (factor correlations) at best would only run about .54 (assuming .5 reliability). It must be remembered that the experiment was not designed to measure the factor, but only to identify it, and identification is inferred from (a) the marker set to load this factor loading it significantly, the bulk of its com-

mon variance being accounted for by this factor and no other, (b) in the case of two markers for one factor from their loadings falling appropriately together on one factor, (c) the fact that no pair of MMPI parcels load the factor more highly than the markers, and, (d) the fact that no seriously competing alternative empirical pattern exists for the identification.

Because doubt could still exist in a complex judgment despite four such sources of evidence (for example, if the marker correlation is quite low, one might argue that it is no more than the correlation of a marker for X with a factor Y decidedly oblique to X), some brief discussion of individual factors is needed. For example, Factor 1 is considered 16 PF Factor A (affectothymia) characterized by outgoing behavior. Though the marker loads on .30, the only other variables on Factor 1 surpassing this, except Parcels 17, Departing from the norm in occupational choice (.41), and 41, Freedom from external control (— .30), are highly consistent with the affectothymic interpretation. For Factor 2 the highest loading, .57, is Factor B (intelligence) on the 16 PF. No other significant loadings for this factor appeared. On Factor 3, the identification is not too satisfactory, because the C factor on the 16 PF (ego strength) loads only .25, and, in this special case, there are some six MMPI parcels which load as well or better: Welsh's repression scale (.37), Psychopathic deviate (false) (.36), Agreement with general external standards (false) (.33), Expression of autonomy (true) (— .32), Fundamental religiosity (true) (— .31), and Denial of masculine occupations (false) (.30). Again, however, the new variables have a meaning fully consistent with ego strength and in no case does the C factor marker on the 16 PF load any other factor comparably with Factor 3. Factor 4 is loaded .33 by the 16 PF dominance marker. Although one other 16 PF variable, L, protension, suspicious, tops this loading with .43, the supporting loading of the MMPI variable Female sexual identification (— .32) strongly suggests a dominance interpretation. It may be further noted that 16 PF Factor L turned up later as a separate factor, principally defined by 16 PF L (.32), Faking to seem good (False) (.29) and Per-

sonal and emotional sensitivity (false) (.46), consistent with the protensia interpretation. Factor 5 is loaded .43 by the surgency marker; the nearest other loading being — .41 by the MMPI parcel, Personal and emotional sensitivity (true). Surgency did not load above .19 on any other factor.

Parenthetically, since readers accustomed to analysis at a scale-correlational rather than factor-analytic level may be suspicious of inference from such comparatively low correlations, a word on method must be inserted. The present study is a first blocking in of a large area, which will, now that specific hypotheses have received tentative support, justify a major research. To get 64 scale measures in 4 hours of available testing time each had to be short to the point where, even had it been a pure factor measure, it would not load more than .77 (assuming a reliability of no more than .6 for the average 10-item scale). We do not get a tidy diagonal of .77 loadings athwart the reference vector structure, but we do get a total structure which will not fit any alternative rearrangement one likes to try. The demonstration of the position of a factor depends on the accuracy of hyperplane determination, which is excellent in the obtained structure, not on the presence of high loaded variables. Except for H and Q_s, to be discussed, the marker loadings are statistically significant beyond $p = .01$ by Harris's (1964) test. The contributions of these markers to the variance within any other factor are trivial and are generally the highest on the factor concerned.

Markers for the super ego, Factor G, the reader will recall, were experimentally split into two. These consistently load .47 and .27 on the same factor, No. 6. With the psychologically related self-sentiment strength factor, Q_s, on the other hand, a real problem arises. As not infrequently happens with versions of Q_s scales, some of its variance actually belongs to the super-ego factor (No. 6 here) and elsewhere, and only in Factor 15 do we find it having variance on a factor not otherwise identifiable. This factor is actually much more clearly marked by MMPI parcels which justify our calling it the general psychoticism factor (mentioned below) and the Q_s loading is trivial. Though no identification

with Q_3 can therefore be made, a hypothesis for further work may be noted that Q_3 , in mixed normal and pathological groups, has some strange relation to psychoticism. Actually, this has been noted before (Cattell, Tatro, & Komlos, 1964) and has given rise to the theory that raised concentration on the approved self pattern (Q_3) is a compensatory device when the ego (C) is psychotically disintegrating.

Factor I, tendermindedness, heavily loaded the eighth vector (.55) supported by the next highest loading (.27) of Female sexual identification.

Autia, M factor, characterized by imaginative behavior, is well marked (.47) and so also is N, Shrewdness. Guilt-Proneness, O factor, has a low loading (.26) on a factor which is loaded at the same time on the obvious reference variable of Edward's MMPI social desirability measure (.42). This is consistent with the conclusion (Cattell, 1957) that anxiety, and particularly the guilt proneness and inadequacy aspects of anxiety contributed by O factor, express themselves by lower social desirability scores. Radicalism, Q_1 , stands out on its own with virtually no MMPI loadings, as would be anticipated from Thurstone's (1947) original definition of this normal factor. The 16 PF self-sufficiency scale, Q_2 , loads a dimension—Factor 14—.54, which is about its usual validity, that is, as a single short scale against the pure factor, and Factor 14 has only trivial MMPI loadings, consistent with the normal associations commonly found for Q_2 . Ergic tension, or tenseness, Q_4 , has a loading of .31 on a factor and is accompanied expectedly, as an anxiety measure, by a negative loading (— .34) on an Edward's social desirability referent in the MMPI.

After these first 16 factors, which, with the two exceptions (H and Q_3) are marked most consistently by 16 PF markers, the remaining factors have very little of 16 PF variance and are clearly defined largely by MMPI variables. Factor 17 is substantially loaded on the Hypochondriasis parcel items 6 and 7 (.50 and .34, respectively), and somewhat less loaded on the Depression, (.16 and .30), Faking bad (.30 and .39), and Hysteria (.39), and may perhaps best be called Hypochondriasis. Factor 18 is recognizable as being probably a

psychopathic deviate factor since it loads the Psychopathic deviate marker (.49) on the MMPI and also to a lesser degree Hypomania (.21 and .21) and Paranoia (.24), which are consistent with the 16 PF profile for psychopaths (high scores on dominance, surgency, and suspicion; Cattell & Eber, 1967).

The third new factor, No. 19, may be interpreted in one or both of two ways. The most obvious interpretation, based on the only two consistently high positive MMPI parcel loadings on it, those of the true and false subsets of Psychasthenia items and substantiated by low social desirability, is that this factor should be labeled Psychasthenia. However, closer examination of the marking variables shows a noticeable split between true and false subsets, for example, those of Social introversion and Mutuality. This latter interpretation is reinforced by the high loading of the Acquiescent response set, which may very well be an alternative for this factor. This finding suggests that, if the labels for these parcels are accurate, psychasthenia may be highly contaminated by an acquiescence response set. Factor 20 has no pathology content as such; its expressions are primarily strong tendency toward nonagreement responses, fundamentalist religion, absence of depression, low dominance, and conservatism. Although the highest marker for this factor was the Negative response set, such an interpretation was not supported by the loadings of the previously mentioned true and false subsets of the MMPI parcels. Hence, Factor 20 is labeled Religiosity with reference to the highest consistent pair of subscale loadings on it: those of Fundamental religiosity. It suggests an assured, resigned, simple faith, and though it occurs in MMPI items it would seem gratuitous from the content and associations to label it as a pathological dimension. Factor 21 plays the usual role of the error factor—the last bit of covariance into which "garbage" gets rotated more or less automatically. Before this, however, comes the fourth of the meaningful pathological factors (No. 15) set aside for discussion because of its curious possible Q_3 association. This has every claim to be considered the "general functional psychosis" factor, hypothesized and set out with preliminary Q -

data evidence and discussion by Cattell in 1957. It loads MMPI parcels for Schizophrenia, Depression, Paranoia, and Hysteria, as well as for the general subjectivity and autism of "anatomy." This has the loss of reality contact and inappropriateness of emotional expression common to diverse psychoses, which have been hypothesized as the hallmarks of the "general psychoticism" concept.

Because of the instances discussed above where factor identification has had to rest—as far as direct loadings are concerned—on values below the .4 or .5 one would expect on psychometric grounds, it has seemed desirable to check by pursuing the independent source of evidence still available. This check on direct validity exists in the calculation of indirect or "circumstantial validity" (Cattell, 1964). Does the factor identified by its internal loadings here as X correlate with other factors as other researches show X to do? This is most completely evaluated by observing its behavior in the second-order domain,

and such a second-stratum structuring would be desirable also to increase our knowledge of the new four pathological factors.

Accordingly the R_o (reference vector correlation matrix) from the above rotations was converted to the R_f for the 21 factors and factored, yielding eight factors by the scree test, which were rotated 12 times to a maximum hyperplane count. All three intermediate-step calculation tables, R , V , and L , are deposited at NAPS as Table F of No. NAPS-00121 (see Footnote 5, and only Table 4, the final simple structure outcome, is presented here. Surveys presently being conducted at Cattell's Laboratory of Personality and Group Analysis, which describe the factors resulting from a dozen studies on second-order personality structure in Q data, indicate that the four largest second-stratum factors here are clearly those usually obtained from personality primaries, namely, Extraversion, Anxiety (closely approximated by Welsh and Dahlstrom's, 1956, estimation, as confirmed by the weighting patterns of Welsh's A scale,

TABLE 4
SECOND-ORDER STRUCTURE OF 16 PF FACTORS FOUND IN MMPI AND MMPI FACTORS
SECOND-STRATUM FACTORS BY NUMBER

First-stratum factors	1 Exvia (Extra- version)	2 Anxiety	3 Cor- tertia	4 Inde- pen- dence	5	6	7	8
1. A Affectothymia	.48	-.20	-.29	.02	.19	.63	-.18	-.11
2. B Intelligence	.06	-.17	-.05	.10	-.02	-.38	.87	.09
3. C Ego Strength	-.17	-.26	.09	.03	-.41	-.22	-.52	-.02
4. E. Dominance	.58	-.00	.08	.05	.58	-.41	.25	-.11
5. F Surgency	.59	.04	-.02	-.13	-.00	.46	-.03	-.08
6. G Super Ego	-.07	-.04	.02	-.22	-.17	.16	.65	.11
7. H Parmia	.53	-.47	.02	-.04	.33	-.02	.07	-.08
8. I Premsia	.03	-.03	-.69	-.03	-.10	-.79	.12	-.13
9. L Protension	.09	.20	.29	.22	.04	-.20	-.47	-.07
10. M Autia	-.04	-.41	-.15	.39	-.23	-.25	.05	.10
11. N Shrewdness	-.05	.01	.13	.04	-.10	.23	.58	.12
12. O Guilt Proneness	.06	.04	.26	-.17	.03	-.62	-.08	.01
13. Q ₁ Radicalism	.08	-.03	-.01	.17	.04	.60	-.53	.48
14. Q ₂ Self-Sufficiency	-.50	.06	.07	.63	-.05	.20	-.00	.05
15. General Psychoticism (Q ₃)	-.00	-.30	-.05	-.00	.30	.95	.04	-.12
16. Q ₄ Ergic Tension	-.07	.19	.10	.07	.31	-.48	-.39	-.06
17. Hypochondria	.03	.46	.01	.01	-.05	.35	.13	.08
18. Psychopathia	-.05	-.03	.00	-.06	.41	.03	.36	.40
19. Psychasthenia	.06	-.14	-.29	-.17	.82	.42	-.02	-.05
20. Religious Resignation	-.57	-.20	-.10	-.20	-.01	.10	.10	-.24
21. Error	-.25	-.09	.04	.03	-.10	.13	.49	-.08

in the top row of Table 1), Inactivity, and Independence. Beyond these, also, as usual, the last two or three factors lack certainty.

From Table 4 one concludes, therefore, that the factors identified internally by marker content as 16 PF factors (in MMPI item space) also occupy the positions relative to one another which factors of this identity would uniquely possess. Regarding next the four new factors, one finds that psychasthenia defines a second order largely to itself (No. 5), as also does general psychosis (No. 6). The psychopathic factor is powerfully affected by some second stratum entity which also favors radicalism. Only the Hypochondriasis^a and Religious Resignation factors have transactions with secondaries already known in the general normal personality field, about half of their variances being due to the general anxiety factor and Introversion, respectively. Whereas the remaining Hypochondriasis variance falls into the pathological Factor 6, Religiosity divides its remainder between Severity (negative Anxiety) and Dependence, both normal second-order factors.

In terms of clinical perspective, what emerges from the present study is that pathological behavior recorded in a broad range of clinical items, such as the MMPI pool, is partly accounted for by general personality source traits common to normals and patients and partly by dimensions which appear with any appreciable variance only in patients. It is appropriate, however, first to pick up and define the latter in mixed samples, as here, in order to get the normal-abnormal structural relations, and also for the practical reason that scales for diagnosis must be designed for diagnosis in populations containing both. However, supplementary refined analyses need to be done on entirely clinical populations, as in the work of Lorr (1955) and Wittenborn (1951).

^a One should be careful not to confuse this hypochondriasis factor, which is a source trait, with the MMPI hypochondriasis surface-trait scale.

REFERENCES

- ADAMS, D. K., & HORN, J. L. Non-overlapping keys for the MMPI scales. *Journal of Consulting Psychology*, 1965, 29, 284.
- CATTELL, R. B. *Description and measurement of personality*. New York: World Book, 1946.
- CATTELL, R. B. *Personality and motivation structure and measurement*. New York: World Book Co., 1957.
- CATTELL, R. B. Evaluating interaction and non-linear relations by factor analysis. *Psychological Reports*, 1960, 7, 69-70.
- CATTELL, R. B. Theory of situational, instrument, second order and refraction factors in personality structure research. *Psychological Bulletin*, 1961, 58, 160-174.
- CATTELL, R. B. Validity and reliability: A proposed more basic set of concepts. *Journal of Educational Psychology*, 1964, 55, 1-22.
- CATTELL, R. B. Evaluating therapy as a total personality change: Theory and available measurements. *American Journal of Psychotherapy*, 1966, 20, 69-88. (a)
- CATTELL, R. B. (Ed.) *Handbook of multivariate experimental psychology*. Chicago: Rand McNally, 1966. (b)
- CATTELL, R. B. The scree test for number of factors. *Multivariate Behavioural Research*, 1966. (c)
- CATTELL, R. B., & BJERSTEDT, A. The structure of depression by factoring Q-data in relation to general personality source traits in normal and pathological subjects. *Scandinavian Journal of Psychology*, 1967, 8, 17-24.
- CATTELL, R. B., & COULTER, M. A. Principles of behavioral taxonomy and the taxonome program. *British Journal of Mathematical and Statistical Psychology*, 1966, 19, 237-269.
- CATTELL, R. B., & DIGMAN, J. M. A theory of the structure of perturbations in observer ratings and questionnaire data in personality research. *Behavioral Science*, 1964, 9, 341-358.
- CATTELL, R. B., & EBER, H. W. *Handbook for the sixteen personality factor questionnaire*. Champaign, Ill.: Institute for Personality & Ability Testing, 1967.
- CATTELL, R. B., & SCHEIER, I. H. Predicting anxiety from clinical symptoms of anxiety. *Psychiatric Quarterly Supplement*, 1961, 35, 114-126.
- CATTELL, R. B., TATRO, D., & KOMLOS, E. The diagnosis and inferred structure of paranoid and non-paranoid schizophrenia, from the 16 PF profile. *Indian Psychological Review*, 1964, 1, 52-61.
- COMREY, A. L. A factor analysis of items on the MMPI Depression Scale. *Educational and Psychological Measurement*, 1957, 17, 578-585. (a)
- COMREY, A. L. A factor analysis of items on the MMPI Hypochondriasis Scale. *Educational and Psychological Measurement*, 1957, 17, 568-577. (b)
- COMREY, A. L. A factor analysis of items on the MMPI Hysteria Scale. *Educational and Psychological Measurement*, 1957, 17, 586-592. (c)
- COMREY, A. L. A factor analysis of items on the MMPI F Scale. *Educational and Psychological Measurement*, 1958, 18, 621-632. (a)
- COMREY, A. L. A factor analysis of items on the MMPI Hypomania Scale. *Educational and Psychological Measurement*, 1958, 18, 313-323. (b)

- COMREY, A. L. A factor analysis of items on the MMPI K Scale. *Educational and Psychological Measurement*, 1958, 18, 633-639. (c)
- COMREY, A. L. A factor analysis of items on the MMPI Paranoia Scale. *Educational and Psychological Measurement*, 1958, 18, 99-107. (d)
- COMREY, A. L. A factor analysis of items on the MMPI Psychasthenia Scale. *Educational and Psychological Measurement*, 1958, 18, 293-300. (e)
- COMREY, A. L. A factor analysis of items on the MMPI Psychopathic Deviate Scale. *Educational and Psychological Measurement*, 1958, 18, 91-98. (f)
- COMREY, A. L. A factor analysis of items on the MMPI Schizophrenia Scale. *Educational and Psychological Measurement*, 1958, 18, 301-311. (g)
- COMREY, A. L., & NENCINI, R. Factors in MMPI responses of Italian students. *Educational and Psychological Measurement*, 1961, 21, 657-662.
- EBER, H. W. Toward oblique simple structure: Maxplane. *Multivariate Behavioural Research*, 1966, 1, 112-125.
- EDWARDS, A. L. *The social desirability variable in personality assessment and research*. New York: Dryden, 1957.
- FINNEY, J. C. Factor structure with the new set of MMPI Scales and the formula correction. *Journal of Clinical Psychology*, 1966, 22, 443-449.
- HARRIS, C. C. Notes concerning the significance of factor pattern coefficients. Department of Educational Psychology, University of Wisconsin, Madison, 1964, 1-4.
- HATHAWAY, S. R., & MCKINLEY, J. C. *Minnesota multiphasic personality inventory manual*. New York: Psychological Corporation, 1951.
- KASSEBAUM, G. G., COUCH, A. S., & SLATER, P. E. The factorial dimensions of the MMPI. *Journal of Consulting Psychology*, 1959, 23, 226-236.
- LAForge, R. A correlation study of two personality tests. *Journal of Consulting Psychology*, 1962, 25, 402-411.
- LINGOES, J. C. MMPI factors of the Harris and Weiner subscales. *Journal of Consulting Psychology*, 1960, 24, 74-83.
- LOVE, M., JENKINS, R. L., & O'CONNOR, J. P. Factors descriptive of psychopathology and behavior of hospitalized psychotics. *Journal of Abnormal and Social Psychology*, 1955, 50, 78-86.
- MEER, H. L., GOCKA, E. F., & HOLLOWAY, H. Social desirability values for California Psychological Inventory items. *Psychological Reports*, 1964, 15, 147-158.
- MESSICK, S., & JACKSON, D. N. Acquiescence and the factorial interpretation of the MMPI. *Psychological Bulletin*, 1961, 58, 299-304.
- SEURE, G. H., & ROGERS, M. S. Note of caution on the factor analysis of the MMPI. *Psychological Bulletin*, 1965, 63, 14-18.
- THURSTONE, L. L. Multiple factor analysis. Chicago: University of Chicago Press, 1947.
- TYLER, F. T. A factor analysis of fifteen MMPI scales. *Journal of Consulting Psychology*, 1951, 15, 451-456.
- WELSH, G. S. A factor study of the MMPI using scales with item overlap eliminated. *American Psychologist*, 1952, 7, 341.
- WELSH, G. S. Factor dimensions A and R. In G. S. Welsh & W. G. Dahlstrom (Eds.), *Basic readings on the MMPI in psychology and medicine*. Minneapolis: University of Minnesota Press, 1956.
- WELSH, G. S., & DAHLSTROM, W. G. (Eds.), *Basic readings on the MMPI in psychology and medicine*. Minneapolis: University of Minnesota Press, 1956.
- WHEELER, W. M., LITTLE, K. B., & LEHNER, G. F. J. The internal structure of the MMPI. *Journal of Consulting Psychology*, 1951, 15, 134-141.
- WIGGINS, J. S. Strategic method and stylistic variance in the MMPI. *Psychological Bulletin*, 1962, 59, 224-242.
- WIGGINS, J. S. Substantive dimensions of self report in the MMPI item pool. *Psychological Monographs*, 1966, 80(22, Whole No. 630).
- WITTENBORN, J. R. Symptom patterns in a group of mental hospital patients. *Journal of Consulting Psychology*, 1951, 15, 290-302.

(Received January 17, 1968)

PERSONAL CHARACTERISTICS OF COLLEGE VOLUNTEERS IN MENTAL HOSPITALS¹

PAUL D. HERSCH

University of Connecticut

JAMES A. KULIK AND KARL E. SCHEIBE²

Wesleyan University

Groups of male and female college students volunteering for summer work in mental hospitals were compared to control students on personality, occupational-interest, and life-history variables. Striking personality characteristics of mental health volunteers were maturity and control, drive for independent achievement, and sensitivity to people and human problems. Vocational interests and life-history data on volunteers were consistent with these personality characteristics. In vocational interests, volunteers were similar to men and women in the independent professions and in professions emphasizing social service or the exercise of language and artistic skills. Life-history data indicated that volunteers were more service-oriented and committed to mental health work than other college students. The role of test-taking attitudes in producing these differences was discussed.

One response to the manpower shortage in the mental health professions has been an increasing use of volunteers in treatment programs for mental patients. The volunteers are most often college students who visit chronic patients on the back wards of mental hospitals. The students not only learn about the field of mental health through such programs, but they also provide needed human contact for mental patients.

Gelineau and Kantor (1964) have reported on the personal characteristics of Harvard and Radcliffe students who make regular visits to a mental hospital during the academic year. These authors emphasized that volunteers were not deviants, but in many ways typical Harvard-Radcliffe citizens. Females, however, were highly overrepresented among volunteers, and the volunteers were distinctive in their occupational interests. The volunteers emphasized service as a source of job satisfaction more than did their peers, and they tended to

prefer mental health careers to careers in business, law, and the natural sciences.

Knapp and Holzberg (1964) have characterized college volunteers in a similar program. In general, these authors concluded that Wesleyan University students who made weekly visits to the Connecticut Valley Hospital were not strikingly differentiated from the general student body. They found some evidence, however, that the companions were more intracceptive, morally concerned, personally compassionate, and introverted than the control group with which they were compared. Volunteers also differed from controls on certain academic variables—major, incidence of disciplinary action, and fraternity membership (Holzberg, Knapp, & Turner, 1967).

The Service Corps program of the State of Connecticut is similar to the programs at Harvard-Radcliffe and at Wesleyan in conceptions and goals. It differs, however, in the amount and intensity of the volunteers' contact with mental patients. The Service Corps students commit themselves to 40 hours a week of interaction with chronic patients for 8-10 summer weeks. The present report focuses on the personal characteristics of these students since it seems possible that the special intensity of their commitment leads to a more distinctive group than those previously studied. The aim of this report is to describe

¹ The research here reported was completed as a part of the Connecticut Service Corps Research Project and was supported by Grant MH02127-03 from the National Institute of Mental Health. The authors are indebted to Richard J. Wiseman, director of the Connecticut Service Corps, whose willing cooperation made the data collection possible.

² Requests for reprints should be sent to Karl E. Scheibe, Department of Psychology, Wesleyan University, Middletown, Connecticut.

these students in terms of personality characteristics, interest patterns, and life-history variables.

METHOD

Subjects

The experimental Ss were 41 male and 110 female students who served in the 1965 and 1966 Connecticut Service Corps (SC). These Ss represent the entire group who completed applications and accepted positions in the SC in these years.

The control groups consisted of 66 females and 76 males enrolled in summer school at four Connecticut colleges. Each member of the control groups was paid \$10 for his participation.

A preliminary comparison of the SC and control groups indicated that the two samples came from similar populations with respect to demographic and sociological variables. Chi-square comparisons between SC and control groups proved to be nonsignificant by conventional standards for the variables of year of college graduation, number of siblings, ordinal position, church attendance, religious attitude, and mother's educational level. Father's educational level was not significantly different for the males, but the educational level of fathers of the SC females

was significantly higher than that of the comparison females ($p < .05$). No significant differences were found by *t*-test comparisons on the variables of age, father's age, mother's age, and family income.

Procedure and Materials

All SC volunteers were administered a battery of tests and questionnaires in group sessions during their first week at the hospitals. The test battery was about 4 hours in length and was completed prior to any patient contact. Members of the control groups were administered an identical test battery at their respective colleges at about the same time.

Measures to be reported cover the areas of personality, occupational interests, and life-history data. As a multidimensional measure of personality, the California Psychological Inventory (Gough, 1957) was selected. Other personality information was obtained by administering the Adjective Check List (ACL; Gough & Heilbrun, 1965), the Rotter (1966) Internal-External (I-E) scale and the Marlowe-Crowne Social Desirability scale (MC-SDS; Crowne & Marlowe, 1964). The Strong Vocational Interest Blank (Strong, 1959) was used as a measure of occupational interests. In addition, five life-history variables derived from a biographical questionnaire were included as indicators of interests and attitudes.

TABLE 1

COMPARISON OF SERVICE CORPS AND CONTROL GROUPS ON THE CALIFORNIA PSYCHOLOGICAL INVENTORY

Class	Scale	Males			Females		
		Service Corps <i>M</i>	Control group <i>M</i>	<i>t</i> test	Service Corps <i>M</i>	Control group <i>M</i>	<i>t</i> test
I	Dominance	54.00	55.61	-.64	52.00	50.48	.91
	Capacity for Status	56.34	55.50	.49	54.77	51.55	1.96
	Sociability	49.29	53.34	-2.06*	51.19	52.14	-.64
	Social Presence	56.93	59.12	-1.09	56.62	53.86	1.54
	Self-Acceptance	57.78	62.12	-2.25*	56.38	58.61	-1.39
	Sense of Well-Being	42.46	42.08	.17	44.44	41.16	1.92
II	Responsibility	50.83	48.97	1.04	50.07	47.81	1.61
	Socialization	47.85	47.67	.09	46.41	46.31	.06
	Self-Control	44.51	40.22	2.23*	44.26	40.22	2.61**
	Tolerance	51.44	48.53	1.56	52.67	46.94	3.42**
	Good Impression	45.66	41.79	2.13*	45.08	41.75	2.33*
	Communality	49.17	50.22	-.60	47.97	48.45	-.34
III	Achievement via Conformity	49.78	47.92	.85	48.07	45.14	1.92
	Achievement via Independence	60.81	56.65	2.24*	59.38	53.28	4.04**
	Intellectual Efficiency	51.93	51.37	.28	53.44	48.25	3.08**
IV	Psychological-Mindedness	57.24	53.26	1.89	55.00	50.95	2.60*
	Flexibility	63.88	58.11	2.38*	61.66	57.56	2.27*
	Femininity	56.10	49.74	3.96**	49.74	52.38	-1.81

* $p < .05$.

** $p < .01$.

RESULTS

Personality Measures

Table 1 lists the means of the 18 scales of the California Psychological Inventory (CPI) for the SC and control groups, together with the *t* ratios indicating the significance of the difference on each scale. Thirteen of the 36 comparisons reach significance at or beyond the .05 level.

Class I scales of the CPI are generally interpreted as measuring social poise, ascendancy, and self-assurance. Scores of the SC males, but not of the SC females, are relatively low on this group of scales. Both male and female volunteers are higher than the control Ss on Class II and Class III scales of the CPI. Class II scales are usually taken to be measures of maturity and control while Class III scales measure achievement potential and intellectual efficiency. On Class IV measures, SC males and females are more psychologically minded and flexible than their peers, but there is a sex difference on the femininity scale. The SC males are clearly more feminine or nurturant than nonvolunteers whereas the female members of the SC appear to be slightly less feminine than their comparison group.

In summary, SC volunteers appear to be mature, tolerant, and controlled; high in need for achievement, especially creative achievement; and psychologically minded and flexible in their thinking. Male volunteers, but not female volunteers, tend to be less socially poised and more nurturant than their non-volunteering peers.

A very similar characterization of the SC females emerges from an analysis of ACL results. SC females are significantly higher on scales reflecting self-control and achievement, and lower than controls on scales of Heterosexuality and Succorance. SC males are higher than controls on the Abasement variable of the ACL.⁸ No differences were found between SC and control for either sex on the I-E or MC-SDS variables.

⁸ Tables giving complete data for each subgroup have been deposited with the American Society for Information Science. Order NAPS Document 00213 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001; remitting \$1.00 for microfiche or \$3.00 for photocopies.

Vocational Interests

In Table 2 are found the means for the male groups on 31 of the 51 occupational scales of the Strong Vocational Interest Blank (SVIB). On these scales, the two groups differ significantly ($p < .05$).

In vocational interests, SC males most resemble men in occupational Groups I, V, VI, and X. Group I occupations are all professions which tend to be practiced independently. The four differentiating scales in Group V—vocational counselor, social worker, school superintendent, and minister—are directly concerned with welfare and service. Group VI and Group X occupations require artistic and verbal-linguistic interests.

The vocational interests of SC males are clearly unlike those of certain occupational

TABLE 2
COMPARISON OF SERVICE CORPS MALES AND CONTROL GROUP ON THE STRONG VOCATIONAL INTEREST BLANK

Group	Occupation	Service Corps M	Control group M	<i>t</i> ratio
I	Artist	41.2	31.0	4.84**
	Psychologist	46.6	37.8	3.82**
	Architect	36.7	29.2	3.34**
	Physician	43.9	37.7	2.39*
	Psychiatrist	45.5	37.2	3.57**
	Veterinarian	13.5	17.3	-1.99*
II	Mathematician	31.2	25.6	2.56*
III	Production manager	21.5	30.1	-4.62*
IV	Farmer	23.7	28.6	-2.33*
	Aviator	23.7	31.5	-3.03**
	Industrial arts teacher	5.0	11.0	-2.29*
	Vocational agriculture teacher	12.7	17.2	-2.02*
	Policeman	21.6	25.7	-2.48*
	Army officer	17.9	27.4	-3.64**
V	Vocational counselor	39.4	35.5	2.06*
	Social worker	44.0	35.6	3.56**
	School superintendent	34.7	28.4	2.57*
	Minister	36.5	25.5	4.35**
VI	Musician	50.5	39.1	5.05**
	Music teacher	42.5	32.2	4.00**
VIII	Senior CPA	27.5	34.2	-3.21**
	Accountant	17.0	24.8	-4.09**
	Office worker	22.0	29.0	-3.96**
	Credit manager	28.1	33.4	-2.64**
	Purchasing agent	13.6	24.4	-5.52**
	Banker	19.7	23.4	-2.26*
	Pharmacist	26.8	30.8	-2.44*
IX	Sales manager	27.1	31.4	-2.18*
X	Advertising man	44.4	39.2	2.94**
	Lawyer	43.4	39.4	4.38**
	Author-journalist	46.2	38.2	4.64**

* $p < .05$.

** $p < .01$.

TABLE 3

COMPARISON OF SERVICE CORPS FEMALES AND CONTROL GROUP ON THE STRONG VOCATIONAL INTEREST BLANK

Group	Occupation	Service Corps M	Control group M	t ratio
II	Artist	36.8	28.8	3.00**
		37.9	30.6	2.61*
III	Social worker	42.5	35.4	2.57*
		36.4	22.0	5.30**
IV	Life insurance salesman	20.5	26.9	-2.14*
V	Buyer	14.0	26.2	-4.54**
	Business education teacher	14.7	25.6	-3.74**
	Stenographer-secretary	27.8	36.5	-3.81**
	Office worker	20.1	31.8	-5.15**
VI	Elementary teacher	25.8	32.8	-2.67**
	Housewife	26.8	35.8	-4.12**
	Dietitian	15.5	22.6	-2.39*
VIII	Physician	30.8	21.8	2.89**

Note.—SVIB scores were not collected from females in 1964. Consequently, this analysis is based on a reduced sample of 68 SC and 20 control females.

* $p < .05$.

** $p < .01$.

groups. In occupational interests, SC males do not resemble members of the technical or skilled trades (Group IV). Furthermore, the SC male does not share the interests of men in business administration or business detail (Group VIII). On seven of the eight scales in Group VIII, the SC are significantly lower than the comparison group.

Table 3 presents comparable data from the SVIB for women. Thirteen of the 29 comparisons are significant at or beyond the .05 level. SC females, like males, are distinctively high on occupations involving verbal-linguistic and artistic interests—artist, author, and music performer. Females are also high on the Group III social service occupations and on the physician scale. On the other hand, the interests of female volunteers are unlike those of women in routine business and teaching posts. The control group is much higher than the SC group on the “nonprofessional” cluster of occupations—office worker, stenographer, and housewife.

In general, male and female SC volunteers have vocational interests in the independent professions; in social service; in careers involving the use of language or the exercise of artistic skill. By temperament, the SC volunteer seems unsuited to business, entrepreneurial, and “nonprofessional” occupations.

Life-History Variables

Five biographical items included on the questionnaire asked students about mental health interests and volunteer activities. Responses to these items were analyzed separately for the male and female groups. Nine of the 10 chi-square comparisons were significant at or beyond the .05 level.

SC volunteers clearly differ from nonvolunteers in their interests in mental health. Volunteers more often plan careers in mental health fields ($p < .01$ for both males and females) and more often choose academic majors in psychology and other mental-health-related fields ($p < .01$ for both males and females). More SC than control males have visited a mental hospital before the SC summer ($p < .01$); female SC, however, do not differ from controls on visits to a mental hospital.

Comparison on biographical questions indicates that volunteer work is more characteristic of SC than control students. More SC members than control students have engaged in both college-related volunteer work ($p < .05$ for males; $p < .01$ for females) and extra-curricular volunteer work ($p < .01$ for both males and females).

In general, the SC volunteers seem more service-oriented and committed to mental health as a profession than other college students.

DISCUSSION

The composite picture of the mental health volunteer that emerges from the CPI is of a student who is mature and controlled, oriented toward independent achievement, and sensitive to people and human problems. The elevation of both male and female SC volunteers on the Good Impression scale of the CPI raises questions, however, about the accuracy of this picture. The possibility must be considered that differences on the Good Impression scale and other scales of the CPI simply reflect attempts by the volunteers to create a favorable impression and to comply with “role demands” of the testing situation.

Some support that the present findings reflect real personological traits rather than test-taking artifact comes from the work of Holzberg, Knapp, and Turner (1967). These re-

searchers compared volunteers and controls on test information available from student files. The test data were obtained under uniform conditions at the time of students' entry into college, before they had expressed any attitudes about volunteer work in mental hospitals. Holzberg and his colleagues found a similar lack of economic motivation, overrepresentation of psychology as an academic major, and orientation to service. In addition, Holzberg refers to the lower incidence of disciplinary action against volunteers during their college careers—confirmation from another domain of data that volunteers are at least as psychologically mature as nonvolunteers.

Consistencies within the present study also lead the present authors to doubt that group differences on the CPI can be attributed completely to "role demands" upon the volunteers. The psychological-mindedness of SC volunteers, for instance, is reflected not only on the CPI but also in their occupational preference for person-oriented careers and in their choice of majors in mental-health-related fields. Drive for independent achievement was expressed both on the personality and interest inventories. The data reported here suggest that participation in volunteer work is not motivated by overconcern with personal problems but rather is partly attributable to a controlled drive for independent achievement and sensitivity to human problems. Such characteristics may have functional significance for volunteer programs. Scheibe (1965) has already noted the low failure rate for these self-selected students, and follow-up question-

naires indicate that the students develop an ever deeper commitment to the mental health professions. Thus, these self-selected students appear capable of responding to the challenge of mental health work and the program appears to crystallize the students' budding interest in the mental health professions.

REFERENCES

- CROWNE, D. P., & MARLOWE, D. *The approval motive: Studies in evaluative dependence*. New York: Wiley, 1964.
- GELINEAU, V. A., & KANTOR, D. Pro-social commitment among college students. *Journal of Social Issues*, 1964, 20, 112-130.
- GOUGH, H. G. *Manual for the California Psychological Inventory*. Palo Alto: Consulting Psychologists Press, 1957.
- GOUGH, H. G., & HEILBRUN, A. B. *The Adjective Check List manual*. Palo Alto: Consulting Psychologists Press, 1964.
- HOLZBERG, J. D., KNAPP, R. H., & TURNER, J. L. College students as companions to the mentally ill. In E. L. Cowen, E. A. Gardner, and M. Zax (Eds.), *Emergent approaches to mental health problems*. New York: Appleton-Century-Crofts, 1967.
- KNAPP, R. H., & HOLZBERG, J. D. Characteristics of college students volunteering for service to mental patients. *Journal of Consulting Psychology*, 1964, 28, 82-85.
- ROTTER, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 1966, 80(1, Whole No. 609).
- SCHEIBE, K. E. College students spend eight weeks in mental hospital: A case report. *Psychotherapy: Theory, research, and practice*, 1965, 2, 117-120.
- STRONG, E. K. *Manual for the Strong Vocational Interest Blanks for Men and Women*. Palo Alto: Consulting Psychologists Press, 1959.

(Received January 25, 1968)

NEED FOR APPROVAL AND POPULARITY: SEX DIFFERENCES IN ELEMENTARY SCHOOL STUDENTS¹

STEVEN R. TULKIN, JOHN P. MULLER, AND LANE K. CONN²

Harvard University

The Children's Social Desirability Scale (CSD) and a sociometric measure were group administered to 192 fifth- and sixth-grade students. Intelligence and achievement scores were also obtained. High need for approval girls (as measured by CSD) were found to be most popular, while high need for approval boys were least popular. Partialing out the effects of intelligence did not significantly alter these sex differences. Sex role theory was used to explain the findings, hypothesizing that behaviors associated with a high need for approval are consistent with the female sex role model but contrary to the male model.

Crowne and Marlowe (1964) presented the view that the tendency to respond in a socially desirable manner is not merely a distortion of test responses, but, more importantly, can be considered as "meaningful problem solving behavior indicative of an individual's approach to being evaluated," and is related to behavior in other situations. Rather than investigate the personality test responses themselves, Crowne and Marlowe postulated a motivational state—the need for approval (n App)—which was reflected in test-taking behavior, and sought its correlates in behavior outside the testing situation. Initial research indicated that high scorers on the Marlowe-Crowne Social Desirability Scale (M-C SDS) tended to conform in an Asch-type experiment (Strickland & Crowne, 1962) and to be more susceptible to manipulation of their attitudes than individuals less motivated toward approval from others (Salman, 1962).

In order to investigate the emergence and development of n App, a Children's Social Desirability Scale (CSD) was constructed

(Crandall, Crandall, & Katkovsky, 1965). Crandall (1966) observed a "striking similarity" between the behaviors characteristic of high n App adults and the personality attributes of high CSD children of high school age as revealed in their scores on the California Psychological Inventory. Specifically, high CSD children showed the same sorts of concern about others' evaluations, the same low self-esteem and lack of self-confidence, and the same conventional, inhibited, suggestible, controlled behaviors and attitudes exhibited by the young adults tested with the M-C SDS in previous studies.

Crandall (1966) also reported that high CSD males and females behaved differently. For the females, significant correlations were reported between CSD score and instigation of physical aggression ($-.40$), verbal aggression ($-.49$), and withdrawal from aggressive attacks ($.33$), while for males the correlations with the same variables were $.14$, $-.08$, and $-.02$. Males with high CSD scores were characterized by infrequent achievement behaviors. They showed significantly fewer attempts at mastery of fine motor skills, spent less time working alone in achievement tasks, made significantly less effort to achieve, and in general tended to be less persistent when they were engaged in achievement activities.

Crandall et al. (1965) also reported large sex differences on the CSD itself. Females had higher scores at every grade tested, and the authors speculated that it "may be more important for American girls to appear socially acceptable than for American boys to do so."

¹This research was partially supported by National Institute of Mental Health Grant HDO 1652 to Lane K. Conn. Computer facilities for the analysis of the data were provided through NSF Grant Number GT2723 to the Laboratory of Social Relations, Harvard University.

²The authors wish to acknowledge the assistance of Walter Gleason, L. Rich, and P. Caffrey in providing the subject population, and to express their gratitude to Kenneth Gergen for his suggestions on interpreting the data and Krayna Tulkin for clerical assistance. Requests for reprints should be sent to Lane K. Conn, Department of Social Relations, Harvard University, Cambridge, Massachusetts 02138.

The present investigation focuses on the relationship between popularity and *n* App: How successful are approval-oriented individuals in achieving social acceptance? Bank³ reported a correlation of $-.33$ between scores on the M-C SDS and likeability ratings among a group of fraternity brothers. High *n* App scores were also associated with a greater expectancy of rejection in TAT stories, and with overestimation of popularity on the sociometric measure. Salman (1965) used female Ss in a similar investigation and found no correlation between M-C SDS scores and popularity ($r = .07$).

Other studies of popularity further define the relationships to be investigated in the present research. Iscoe and Carden (1961) administered the Witkin Embedded Figures Test, the Children's Manifest Anxiety Scale (CAS), and a three-question peer rating index to a sixth-grade class. The Witkin Test is relevant to the present discussion in light of Rosenfeld's (1967) report that high *n* App Ss were field independent. Iscoe and Carden found a significant rank-order correlation of $.57$ ($p < .05$) between sociometric status and field dependence for girls, in contrast to a significant negative correlation of $-.51$ ($p < .05$) for males. In addition, a correlation of $-.60$ ($p < .01$) was obtained between CAS and field dependency for females indicating that field independent girls tended to be more anxious than field dependent ones. For males, the correlation was positive but not significant. The authors explained their results by referring to culturally learned sex-role behavior:

Boys are expected to be somewhat aggressive, direct, and analytic while girls are taught a more submissive and conforming 'ladylike' type of behavior. The girl who identifies with this role gains acceptance . . . (while) an analytic field independent mode of perceiving results in less popularity and more anxiety. The reversal of this relationship in boys perhaps emphasizes the cultural rewards for their exhibiting initiative [Iscoe & Carden, 1961, p. 84].

Similarly Jones, Jones, and Gergen (1963) found that male Ss negatively evaluated male conformers and, in fact, gave the highest ratings to those who did not conform even when under strong pressures to do so. Females, on

the other hand, showed a general tendency to give more positive evaluations to males who conformed.

The purpose of the present study is to further investigate the relationship between approval-oriented behavior and both actual and perceived popularity. It was hypothesized that high *n* App males would be unpopular while the negative relationship between *n* App and popularity would not hold up for the females.

METHOD

Subjects

The Ss for the present study were 192 fifth- and sixth-grade students from a suburban Boston school system. The sample comprised the entire fifth and sixth grades in two schools. There were an equal number of fifth and sixth graders and an equal number of males and females. An estimate of each family's social class was determined by the occupation of the student's father using the Hollingshead (1957) Index of Social Position. Each of the seven occupational levels except for the lowest (unskilled and unemployed) was well-represented.⁴

Measures

The Children's Social Desirability Scale (CSD) Intermediate Level, was group administered to each of seven classes with the teachers present. Previous inquiries had shown that students had no difficulty with reading or understanding the questions. Intelligence and achievement test scores were taken from the students' record cards. Intelligence was measured by the Otis Quick Scoring Test and achievement by the Stanford Achievement Test.

A modified sociogram was administered on which the students first listed names of three classmates for whom they would vote on each of five questions: president of the class, representative to a television quiz show, person played with the most, person liked best, and person liked least. These lists were collected and a second sheet handed out on which each student was asked to guess how many votes he or she would receive on each of the questions. Four variables were created for each question: the actual vote received, the estimated vote, the absolute discrepancy between the vote and the estimate, and a relative discrepancy between the estimate and the vote (hereafter termed "base") which indicated whether the estimate was greater or less than the actual vote. An estimate of the independence of the five sociometric questions was obtained by correlating the number of votes which each student received on each of the five questions. These correlations are presented in Table 1. While some of the questions share a considerable amount of common variance ("president"

³ S. P. Bank, unpublished manuscript, 1963. Cited by Crowne and Marlowe (1964).

⁴ Level 1, 32; Level 2, 17; Level 3, 33; Level 4, 37; Level 5, 48; Level 6, 23; and Level 7, 2.

TABLE 1
INTERCORRELATIONS OF VARIOUS
SOCIOMETRIC MEASURES

Measure	Quiz	Play	Likes best	Likes least
President	.854	.468	.569	-.325
Quiz		.382	.439	-.259
Play			.783	-.393
Likes best				-.384

Note.—All correlations are significant at $p > .01$.

and "quiz," for example), many of the correlations reflect a moderate amount of independence. It appears that two clusters emerge: one dealing with task competence ("president" and "quiz") and the other reflecting a more social or affective popularity ("play" and "likes best").

RESULTS

Table 2 presents descriptive statistics on the CSD for the present sample and that reported by Crandall, Crandall, and Katkovsky (1965). The results are quite similar, except for the sixth-grade males where the difference is significant at the .05 level.

Table 3 presents the correlations of major variables with the scores on the CSD for males and females. The significant differences between the correlations of the males and females indicate that high n App is related to different clusters of behaviors for the two sex groups. For females, high n App is significantly correlated with higher peer ratings on all of the four "positive" questions, and also to higher estimates of the number of votes that would be received. High n App females

TABLE 2

NORMATIVE DATA ON THE CSD FOR PRESENT
SAMPLE AND CRANDALL, CRANDALL,
AND KATOVSKY, 1965

Grade	Present sample			Crandall et al., 1965		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
5						
Females	45	24.47	7.09	53	23.00	9.33
Males	43	20.63	8.51	53	21.75	8.84
6						
Females	46	23.41	8.73	73	22.12	8.82
Males	49	20.67	7.02	93	17.68	8.33

also tend to have higher discrepancies between actual and predicted vote, and the correlations with "base" variables show that the females tend to overestimate. For males, on the other hand, high n App is associated with fewer votes on all questions, although the correlations are significant only for the "friendship" categories ("play with" and "likes best"). Also in contrast to the females, high CSD males give lower estimates of the number of votes they will receive (although only significant for the vote for president). Their self-deprecating tendency is confirmed, however, by the significant correlation of "base least" which shows that the high CSD

TABLE 3

PRODUCT-MOMENT CORRELATIONS OF CSD
WITH OTHER VARIABLES FOR
MALES AND FEMALES

Variable	Males	Females	Significance of differences between correlations
IQ	-.268**	-.067	
Achievement	-.317**	.000	
Vote for			
President	-.129	.294**	.01
Quiz	-.078	.254*	.05
Play	-.286**	.264*	.01
Likes best	-.218*	.350**	.01
Likes least	-.058	-.146	
Estimate for			
President	-.210*	.328**	.01
Quiz	-.127	.349**	.01
Play	-.116	.391**	.01
Likes best	-.098	.329**	.01
Likes least	.171	-.245*	.01
Discrepancy for			
President	-.157	.285**	.01
Quiz	.024	.329**	.05
Play	-.060	.181	
Likes best	-.102	.237*	.05
Likes least	.194	-.164	.05
Base for			
President	-.109	.135	
Quiz	-.066	.173	
Play	.021	.306**	.05
Likes best	.035	.137	
Likes least	.227*	-.112	.05

* $p < .05$; two-tailed test.
** $p < .01$; two-tailed test.

TABLE 4

PARTIAL CORRELATIONS BETWEEN CSD AND
PEER RATINGS HOLDING INTELLIGENCE
CONSTANT

Variable	Males	Females	Significance of differ- ences be- tween cor- relations
Vote for			
President	-.024	.337**	.05
Quiz	.051	.314**	.05
Play	-.263**	.272**	.01
Likes best	-.195*	.365**	.01
Likes least	-.106	-.161	
Estimate for			
President	-.146	.476**	.01
Quiz	-.026	.432**	.01
Play	-.158	.399**	.01
Likes best	-.111	.224*	.01
Likes least	-.166	-.044	
Discrepancy for			
President	-.136	.252*	.05
Quiz	.103	.391**	.05
Play	-.110	.221*	.05
Likes best	-.130	.204*	.05
Likes least	.179	.075	

* $p < .05$; two-tailed test.

** $p < .01$; two-tailed test.

males overestimate the number of votes they think they will receive in the "likes least" category ($p < .05$).

Table 3 also reveals a significant correlation between intelligence and CSD scores for the males. Since intelligence was significantly related to several of the sociometric ratings, the possibility existed that intelligence, rather than *n* App, was responsible for the correlations presented above. To investigate this possibility, partial correlations were computed between CSD and the sociometric measures with the effects of intelligence partialled out. Table 4 presents these partial correlations and demonstrates that, even with intelligence held constant, there are significant differences between the sexes on the relationship between CSD and the sociometric ratings.

DISCUSSION

The present study suggests that high *n* App has an opposite relationship with popularity for male and female fifth- and sixth-

grade students. It therefore seems necessary to specify further the explanations that have been offered for the lack of popularity of the high *n* App individuals, and to suggest that these hypotheses are valid only for males. Marlowe and Gergen (in press), for example, suggested that the high *n* App person works too hard at winning acceptance, is seen as unrealistically accommodating, and, as a consequence, is disliked. Similarly, Barthel (1963) noted that high *n* App males appear to be cautious in their relationships with others, over-dependent upon the guidance of those around them, and eager to create favorable, even if distorted, impressions. Barthel submitted that such behavior, although tolerated, might be perceived as "phony" by persons in close contact with the high *n* App individuals. The present research suggests that these arguments are applicable only to males, while for females this type of behavior tends to be associated with greater popularity.

We might hypothesize that low *n* App females and high *n* App males violate their cultural role patterns and are therefore rejected by their peers. This hypothesis seems to be supported by descriptions of high CSD males as not evidencing achievement behaviors (Crandall et al., 1965) and as significantly more often imitating the play activities of other children (Crandall, 1966). Such behaviors appear to violate sex-role norms for American males. Similarly, low CSD females are described by Crandall (1966) as more frequently "initiating aggression of both physical and verbal nature" and as less frequently withdrawing from attacks. This behavior seems quite incompatible with the feminine sex-role model.

Further clarification is offered by Hilkevitch (1960), who investigated compatibility of personalities among male and female friendship groups. He found that among boys there were significantly more complementary traits than common traits, while for girls, friendship was "based more on similarity than difference of attributes." Girls tended to befriend other girls if they had a lot in common, while males seemed to prefer friendships with others who showed individuality and could complement each other, rather than be identical.

Tables 3 and 4 reported significant sex differences in correlations between CSD and the estimate and discrepancy variables. For the females, this suggests that although those with higher CSD scores received more votes, they were deficient in role taking; that is, they were less accurate in their perceptions of how they were viewed by their peers. The high n App males, on the other hand, were more accurate in their perceptions. We might hypothesize that approval-oriented females orient themselves to their peers as a reference group. On the other hand, higher CSD males, who were less discrepant in their estimations, may be oriented to an adult reference group, and through interactions with adults have encountered more complex role-taking capabilities.

The CSD may have greater predictive utility when further research is undertaken toward understanding how the need for approval is sought. The fact that a person may have many reference groups toward which he orients himself differentially implies that the n App construct would have greater predictive validity if reference group norms were taken into account.

REFERENCES

- BARTHEL, C. E. The effects of the approval motive, generalized expectancy, and situational cues upon goal setting and social defensiveness. Unpublished doctoral dissertation, Ohio State University, 1963.
- CRANDALL, V. C. Personality characteristics and social achievement behaviors associated with Children's Social Desirability response tendencies. *Journal of Personality and Social Psychology*, 1966, 4, 477-486.
- CRANDALL, V. C., CRANDALL, V. J., & KATKOVSKY, W. A children's social desirability questionnaire. *Journal of Consulting Psychology*, 1965, 29, 27-36.
- CROWNE, D. P., & MARLOWE, D. *The approval motive: Studies in evaluative dependence*. New York: Wiley, 1964.
- HILKEVITCH, R. R. Social interactional processes: A quantitative study. *Psychological Reports*, 1960, 7, 195-201.
- HOLLINGSHEAD, A. *The two-factor index of social position*. New Haven: Author, 1957.
- ISCOE, I., & CARDEN, J. A. Field dependency, manifest anxiety, and sociometric status in children. *Journal of Consulting Psychology*, 1961, 25, 184.
- JONES, E. E., JONES, R. G., & GERGEN, K. G. Some conditions affecting the evaluation of a conformist. *Journal of Personality*, 1963, 31, 270-288.
- MARLOWE, D., & GERGEN, K. G. Personality and social interaction. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology*. (2nd ed.) Cambridge: Addison Wesley, in press.
- ROSENFELD, J. M. Some perceptual and cognitive correlates of strong approval motivation. *Journal of Consulting Psychology*, 1967, 31, 507-512.
- SALMAN, A. R. The need for approval, improvisation and attitude change. Unpublished master's thesis, Ohio State University, 1962.
- SALMAN, A. R. The development and validation of a projective measure of the need for approval. Unpublished doctoral dissertation, Ohio State University, 1965.
- STRICKLAND, B., & CROWNE, D. P. Conformity under conditions of simulated pressure as a function of the need for social approval. *Journal of Social Psychology*, 1962, 58, 171-181.

(Received February 2, 1968)

PATIENT-THERAPIST INTERDEPENDENCE: COGNITIVE AND BEHAVIORAL¹

PETER S. HOUTS²

Pennsylvania State University College of Medicine

SHIRLEY MacINTOSH

Stanford University School of Medicine

AND RUDOLF H. MOOS

Stanford University School of Medicine

Each of seven patients saw each of four therapists twice. After each session, therapists rated how they thought the patient and therapists should behave during psychotherapy. In addition, these patient and therapist behaviors were independently rated from tapes of the meetings. A balanced 4×7 design was utilized. Behavior rating data indicate that therapists are more influenced by patients than patients are by therapists. Comparisons of questionnaire and behavior rating data indicate that differences among therapists are more evident in how therapists think than in how they or patients behave. Implications for understanding the different roles of the psychotherapist are discussed.

While the importance of viewing the psychotherapeutic interview as an interdependent system of patient and therapist behaviors has often been asserted, there has been little research explicating the degree and types of such interdependence in psychotherapy sessions. A research design where each of the several patients meets with each of several therapists is one technique for quantitatively studying such interdependence. In such a design, patients can be represented by rows and therapists by columns with one cell for each patient-therapist combination. This design allows data analysis using a two-way analysis of variance. The degree of patient-therapist interdependence can then be assessed by the percentage of variance accounted for by the patient main effects and interactions when studying therapist behaviors, and by the percentage of variance accounted for by the therapist main effects and interactions when studying patient

behaviors. These percentages of variance will indicate the degree to which differences between patients or unique patient-therapist combinations influence therapist behaviors, and the degree to which differences between therapists or unique therapist-patient combinations influence patient behaviors. Conversely, a lack of patient-therapist interdependence can be shown by the size of the therapist main effects when studying therapist behaviors and of patient main effects when studying patient behaviors. The percentage of variance accounted for by these main effects will indicate the degree to which there are consistent differences between patients or between therapists, independent of the person with whom they are meeting.

Two earlier studies utilizing this design have investigated the nature and degree of patient-therapist interdependence. Van der Veen (1965), working with chronic, hospitalized schizophrenics, rated problem expression and immediacy of experiencing for patients and congruence and accurate empathy for therapists. He reported significant patient main effects, therapist main effects, and interactions for both patient and therapist behaviors. He found patient main effects are the strongest of the three factors in patient behaviors, and therapist main effects the strongest factor in therapist behaviors. Moos and Clemes (1967), working with an outpatient clinic population, measured for both patients

¹ This research was supported, in part, by National Institute of Mental Health Grant MH8304. The authors wish to express their appreciation to Eleanor Levine, Phyllis Noble, Karl Schonborn, and Robert Harris for their help as behavior raters, to Helena Kraemer for statistical consultation and to the therapists and patients who participated in this research.

² The collection and analysis of the data was accomplished while this author was associated with Stanford University School of Medicine. Requests for reprints should be sent to Peter S. Houts, Pennsylvania State University, The Milton S. Hershey Medical Center, Hershey, Pennsylvania 17033.

and therapists the total number of words spoken, the percentage of feeling words, the percentage of action words, and the number of reinforcements. They also reported a number of significant patient and therapist main effects as well as interactions. However, in contrast to Van der Veen, they reported that, with the exception of "number of reinforcements," patient main effects accounted for a larger percentage of the variance than did therapist main effects in both therapist and patient behaviors. For most of the behaviors measured, they concluded that patients influence therapists more than therapists influence patients. Methodologically, the Moos and Clemes findings are somewhat stronger than Van der Veen's because Van der Veen's behavior ratings involved subjective estimates by raters who rated all meetings. Thus, raters' expectations about consistent differences between therapists or patients could contribute to the size of either patient or therapist main effects. Moos and Clemes' variables, on the other hand, required very little rater inference and therefore minimized the possibility of rater bias contributing to main effects.

Utilizing the same design as the Van der Veen and Moos and Clemes studies, the present study adds an additional dimension, namely how therapists think patients and therapists should act during the psychotherapeutic interview as well as the behavior ratings which were parallel to the questionnaire measures. The purpose of this study was to compare the percentages of variance accounted for by patients, therapists, and interaction in questionnaires and behavior ratings and thus study the magnitude of patient and therapist interdependence in both types of data.

METHOD

This research was carried out in a clinic where patients routinely rotated for 25-minute therapy sessions among a resident therapist pool. This was the same rotating patient clinic (RPC) used and described by Moos and Clemes. The present sample included seven female patients representative of the RPC population and four male psychiatric residents; two in their second and two in their third years of residency. Each of the seven patients met twice with each of the four therapists for a total of 56 psychotherapeutic interviews, each of approximately 25-minutes duration. Each interview was tape recorded. Following each interview, the therapist rated on a

4-point scale (a) what he thought each patient should do during therapy sessions to accomplish his goals and (b) what he thought the RPC therapists should do during therapy sessions to help the patient accomplish these goals. Questionnaire content was originally obtained through interviews with the RPC patients and therapists. Items which were particularly appropriate for that clinic setting were chosen.

In addition to these postmeeting questionnaires, two raters independently rated the 56 tape recordings using the same patient and therapist behaviors which were in the questionnaire. For example, the therapists' questionnaire asked whether the patient should tell intimate and private details of her life and the behavior raters scored the degree to which the patient did tell intimate and private details of her life during that interview.

RESULTS

Correlations between the two behavior raters range between .37 and .89 on the 20 behavior ratings, with a median of .52. The two raters' scores were summed on each item and these summed scores were used in the analyses. While many of these reliabilities are not high, it will be shown later that our findings are independent of item reliabilities.

Therapist questionnaire scores were transformed into standard scores for each therapist separately, in order to minimize the effects of general response set on therapists' answers. For example, for each therapist, all of his answers to questions about how the therapist should act were combined. Since each therapist filled out the questionnaire 14 times, and each time rated seven items on how therapists should act, the mean and standard deviation of 98 responses were used in transforming the scores to standard scores. These transformed scores can be interpreted as the relative emphasis which each therapist gave to each item. The same procedure was followed with the questionnaire responses to patient behaviors as well as with the independent behavior ratings of therapist and patient behaviors from the taped sessions.

Each of the 20 items in the therapist's questionnaire and each of the 20 ratings of behavior were analyzed with analyses of variance. A two-way design was utilized, four therapists by seven patients, with two replicates per cell.

Table 1 reports results from therapists' questionnaires and independent behavior rat-

TABLE 1

PATIENT BEHAVIORS: PERCENTAGE OF VARIANCE ACCOUNTED FOR BY EACH SOURCE

Patient behaviors	Therapist questionnaires				Behavior ratings			
	Patient	Therapist	Interaction	Within	Patient	Therapist	Interaction	Within
1. Be businesslike toward doctor	6	13 ^a	0	81	54 ^b	0	0	46
2. Tell doctor intimate details of her life	21 ^a	1	6	72	27 ^b	10 ^a	0	63
3. Tell what happened since clinic visit	4	29 ^b	0	67	0	0	0	100
4. Try to talk about what doctor thinks is important	4	53 ^b	4	39	19 ^a	0	0	81
5. Tell about important people in her life	8	13 ^b	32 ^a	47	26 ^b	11 ^a	1	62
6. Talk about her feelings toward doctor	1	66 ^b	11 ^a	22	13	9	0	78
7. Ask the doctor to tell her what to do	3	15 ^a	4	78	21 ^a	7	0	72
8. Tell doctor what other doctors have told her	0	49 ^b	0	51	30 ^b	10 ^a	0	60
9. Summarize her problems for each new doctor	0	2	1	97	2	4	0	94
10. Tell about her physical condition	12 ^a	23 ^b	17	48	80 ^b	0	1	19
11. Let doctor decide what they should talk about	7	13 ^a	0	80	39 ^b	0	24 ^a	37
12. Be logical and unemotional in describing her problems	2	13 ^a	1	84	71 ^b	0	0	29
13. Tell how she gets along with other people	15 ^a	18 ^b	15	52	47 ^b	12 ^b	0	41

^a F value significant at $p < .05$.^b F value significant at $p < .01$.

ings for the patient behaviors in terms of percentage of variance accounted for by each source of variance. Looking first at the therapist questionnaire results, there are more statistically significant therapist effects than either patient main effects or interactions. The percentages of variance accounted for by patient, therapist, and interaction effects was compared using the Friedman nonparametric analysis of variance (Siegel, 1956). The Friedman test is more appropriate for these data than the parametric analysis of variance because the assumptions of equal variance and independence required by the parametric tests cannot be met. The results indicate that the patient, therapist, and interaction percentages of variance do differ significantly ($\chi^2_r = 11.6$, $p < .01$).

Comparison of the average percentage of variance accounted for by the different effects, using the nonparametric Wilcoxon test (Siegel, 1956), shows that therapist main effects

account for more variance than either patient main effects ($z = 2.55$, $p < .01$) or interactions ($z = 2.34$, $p < .05$), and that the patient main effects and interactions do not differ significantly from each other ($z = .82$).

For the behavior ratings on patients, there are more statistically significant patient main effects than either therapist main effects or interactions. A comparison of average percentage of variance accounted for by each effect shows that average patient, therapist, and interaction effects are significantly different ($\chi^2_r = 15.3$, $p < .001$), and that the patient main effects are, on the average, larger than the therapist main effects ($z = 2.98$, $p < .01$), or interactions ($z = 3.18$, $p < .01$). The therapist effects do not differ significantly from the interactions ($z = 1.41$). The therapist main effects are the strongest in the questionnaire results while the patient main effects are strongest in the behavior ratings.

It is also possible to compare directly the

main effects and interactions in therapist questionnaire and independent behavior ratings by computing from their F values an estimate of the noncentrality parameter of the noncentral F distribution (Scheffé, 1959). Using these estimates, we can compare the F statistics for patient main effects, therapist main effects, and interactions between the questionnaire and behavior ratings. These comparisons show that the average patient main effects are larger in the behavior ratings than in the questionnaires ($z = 3.04$, $p < .01$) while the average therapist main effects are larger in the questionnaire results than with the behavior ratings ($z = 2.76$, $p < .01$). There is no statistically significant difference between the two interaction values ($z = 1.89$).

In summary, Table 1 indicates that differences among therapists are very important in how therapists think patients should act, while differences among patients are more important in how patients do act in psychotherapy sessions. It is interesting to note at this point that the four behavior ratings which do show significant therapist main effects (2, 5, 8, 13) all concern the content of what the patients talk about and that none of the ratings about the manner in which the patients act towards the therapist (1, 4, 7, 11, 12) show significant therapist main effects. There is some evidence that therapist differ-

ences influence the content of what patients discuss more than the manner in which patients present that content. Finally, we can note that interactions are not large in either set of data.

Table 2 reports results from therapists' questionnaires and independent behavior ratings for therapist behaviors. The questionnaire data about desired therapist action show approximately an equal number of statistically significant patient main effects and therapist main effects with relatively few statistically significant interactions. Again, using Friedman's nonparametric analysis of variance to compare the average percentage of variance accounted for by the different effects, the findings are that the three effects do not differ significantly ($\chi^2_r = 3.71$).

The independent behavior ratings show evidence only for patient main effects and no evidence of statistically significant therapist main effects. The Friedman test shows that the three effects differ significantly from each other ($\chi^2_r = 8.86$, $p < .02$), and further analysis utilizing the Wilcoxon test shows that the patient main effects account for more variance, on the average, than do therapist main effects ($z = 2.37$, $p < .05$) or interactions ($z = 2.20$, $p < .05$), while the therapist main effects do not differ from the interactions ($z = 1.75$). The transformation of the F values to a non-

TABLE 2
THERAPIST BEHAVIORS: PERCENTAGE OF VARIANCE ACCOUNTED FOR BY EACH SOURCE

Therapist behaviors	Therapist questionnaires				Behavior ratings			
	Patient	Therapist	Interaction	Within	Patient	Therapist	Interaction	Within
1. Suggest new ways of looking at her problems	11 ^b	53 ^b	4	32	10	2	12	76
2. Tell her what to do about her problems	20 ^b	16 ^b	2	62	51 ^b	0	5	44
3. Help her to understand why her personality is the way it is	28 ^b	21 ^b	3	48	33 ^b	0	2	65
4. Speak to other people with whom she has difficulties	24 ^b	27 ^b	26 ^b	23	4	1	0	95
5. Mostly listen to what she has to say	15 ^b	48 ^b	5	32	28 ^b	0	0	72
6. Encourage and support her in what she wants to do	5	22 ^b	6	67	13	0	0	87
7. Give her medicine	50 ^b	6 ^a	16 ^a	28	54 ^b	0	7	39

^a F value significant at $p < .05$.

^b F value significant at $p < .01$.

central parameter estimate was again used to compare directly the questionnaire and behavior ratings. The results show no statistically significant differences between the patient main effects in the two measures ($z = 1.18$), nor between the interactions in the two measures ($z = 1.35$), but do indicate that the therapist main effects are stronger in the therapist questionnaires than they are in the behavior ratings ($z = 2.37, p < .05$). In conclusion, patient main effects are strong both in how therapists think therapists should act and also in how therapists do act. However, therapist main effects are more evident in the thinking of therapists than in their independently rated behavior.

Returning now to the question of whether our findings were influenced by the reliability of the behavior ratings, two points should be made. First, rater reliabilities would influence the residual error terms in our analyses of variance, not the relative sizes of main effects and interactions with which we were concerned. Second, by transforming the F values to a noncentral parameter estimate, we were able to compare the questionnaire and behavior rating results independent of any residual error-term differences. The comparisons of main effects and interactions reported above are independent of rater reliabilities.

Since rater bias may have contributed to the main effects observed in the behavior ratings of both patient and therapist behaviors, four additional raters rescored first meeting tapes from four patients randomly chosen from the seven study patients. These 16 tapes were assigned, four each, to the four new raters in such a way that each rater scored each patient-therapist combination only once. This procedure eliminated the possibility that main effects would be due to a rater imposing a constant stereotype for how a therapist or a patient behaved from one interview to the next. A 4×4 matrix with one replicate per cell was used to analyze both these ratings and ratings of the same meetings from the main study. The F values for each behavior rated were transformed to an F noncentrality parameter estimate and the transformed values for the two sets of ratings were compared. The average main effect and interaction values do not differ significantly between the

two sets of ratings for either patient or therapist behaviors. There is no evidence that rater bias contributed substantially to our main effect findings.

DISCUSSION

Our behavior rating findings, in common with Moos and Clemes', indicate that therapists are more influenced by patients than patients are by therapists. This is indicated by the strong patient main effects and relatively weak therapist main effects in ratings of both patient and therapist behaviors. These findings differ somewhat from those of Van der Veen who reported strong therapist main effects in his measures of therapist behaviors. These differences may be due to differences in the behaviors rated or in the kinds of Ss utilized in the two studies. This study also indicates that one area of therapist influence on patient behaviors appears to be in the content of what patients discuss. Interactions play a negligible role in all of the behavior ratings, a finding consonant with both Van der Veen and Moos and Clemes' results.

The therapist questionnaire findings show an important contrast to those with behavior ratings. While there is evidence for patient differences influencing questionnaire responses, there is also strong evidence for the influence of therapist differences, independent of the patient they are seeing. Interactions again play a negligible role in these findings.

Direct comparison of the size of the main effects and interactions between the questionnaires and behavior ratings, using a transformation of the F values to a noncentral parameter estimate, indicates that differences between therapists carry more weight in the questionnaires than they do in the behavior ratings for both therapist and patient behaviors. This indicates discrepancy between what therapists think should happen in psychotherapy and what does happen and that, in part, these discrepancies are due to the fact that differences between therapists are more important in the thinking of therapists than they are in the behavior of either therapists or patients.

In discussing the implication of these findings, it is important to consider their generality. Clearly, we had a limited sample of

both therapists and patients and, even more important, we were studying an unusual kind of therapy where patients expected to see a different therapist each week. It may be, for example, that long-term intensive psychotherapy provides greater opportunity for consistent differences between therapists (i.e., therapist main effects) to become evident in therapist behaviors. Because our research design requires this somewhat unusual pattern of patient-therapist contacts, it would be difficult to replicate these findings in other types of psychotherapy settings.

While recognizing these limitations, we can also note that our results are consistent with findings reported in other settings and that they can suggest hypotheses to be investigated in other settings as well. For example, the finding that differences in how therapists think about their behavior are not evident in ratings of their behavior in therapy sessions is consistent with Fiedler's (1950) finding that therapists from different theoretical persuasions do not differ in how they behave in therapy sessions. Since many of our questions dealt with issues independent of theoretical schools, this suggests that Fiedler's findings may be just one example of a wide range of discrepancies between therapists' styles in thinking about psychotherapy and what, in fact, transpires in psychotherapy.

One source of these discrepancies may be that the therapist role includes, within it, different expected behaviors in different settings. Two such settings are the psychotherapy interview and the rehash and discussion sessions with fellow psychotherapists. In the psychotherapy interview, the therapist is expected to be sensitive and responsive to patient differences. He learns, both through observing experienced therapists and through being observed himself, that the good therapist is constantly alert to what the patient is feeling and communicating. Our behavior rating data showed that this is, in fact, what occurred in the sessions we studied. In the rehash and discussion sessions, on the other hand, therapists must learn a somewhat different set of skills in dealing with their peers and supervisors. In these sessions, which are particu-

larly prevalent during residency training such as the therapists in our study were undergoing, they must constantly explain their actions to peers and supervisors. This pressure of public explanation may encourage psychotherapists to develop a consistent framework with which they can quickly organize an explanation for their behaviors. As a therapist finds one framework useful for such presentation, he may repeat it until it becomes his style of thinking about therapy. Since these explanations are, for the most part, *ex post facto*, they need not correspond to what in fact occurred in the sessions. What is more important is that they communicate to the audience of fellow professionals that the speaker has a sophisticated grasp of the concepts and rationale of psychotherapy.

In conclusion, our study indicates that in this brief contact clinic setting and for the behaviors we measured, differences between patients are a major determinant of both patient and therapist behaviors. Consistent differences between therapists, on the other hand, are more evident in how therapists think about patient and therapist strategies to achieve psychotherapeutic goals than they are in either therapist or patient behaviors during therapy interviews. We have suggested that these discrepancies between behavior ratings and therapists' questionnaire responses may be due to differences in role expectations for psychotherapists when they are with patients and when they are with professional colleagues.

REFERENCES

- FIEDLER, F. A comparison of therapeutic relationships in psychoanalytic, nondirective, and Adlerian therapy. *Journal of Consulting Psychology*, 1950, 14, 436-445.
- MOOS, R. H., & CLEMES, S. R. A multivariate study of the patient-therapist system. *Journal of Consulting Psychology*, 1967, 31, 119-130.
- SCHEFFÉ, H. *The analysis of variance*. New York: Wiley, 1959.
- SIEGEL, S. *Nonparametric statistics*. New York: McGraw-Hill, 1956.
- VAN DER VEEN, F. Effects of the therapist and the patient on each other's therapeutic behavior. *Journal of Consulting Psychology*, 1965, 29, 19-26.

(Received February 23, 1968)

SUICIDE AND AGE: A SUICIDE NOTE ANALYSIS

ALLEN R. DARBONNE¹

Veterans Administration Center, Los Angeles

The present study tested the hypothesis that suicidal ideation varies with age. 259 suicide notes were analysed. This represented all of the suicide notes written by individuals who committed suicide during a 3-year period who resided in the 34 moderately advantaged natural communities of Los Angeles County. The method of analysis involved categorizing each note according to a scheme of five general categories: (a) Addressee, (b) Reasons Stated, (c) Affect Indicated, (d) Specific Content, and (e) General Focus of Note. The notes were divided into four age groups—20-39, 40-49, 50-59, and 60 and over, and compared with reference to the five general categories. Differences between age groups were found significant at the .05 level for Categories b and e. Specific items within each category were discussed when such items were found statistically more or less often in the notes of one age group than in the notes of the other age groups. Suggestions for treatment and prevention of suicide were offered.

Despite effort and theory which began long before Freud and Durkheim, one of the most pressing questions of society remains unanswered. Why does man so often choose to take his own life? Each year in the United States at least 25,000 persons make this choice, resulting in suicide ranking among the top 10 causes of adult deaths.

There seems to be a growing agreement among current investigators of suicide that present theories do not adequately handle the data. A need has been expressed for an amplification of current conceptual systems to better account for the many variables related to this phenomenon. Among these variables is the factor of age. Statistics clearly indicate that suicide is not restricted to any specific age group. Farberow and Shneidman (1957) pointed out in a theoretical review that there has been no differentiation in suicide theories for age of the suicidal person and indicated the need for this on the basis of their researches.

An understanding of the role of age in suicide would seem to be a desirable early step

in the development of an adequate theory. Suicide is frequently viewed as a result of a combination of the individual's psychic drives and environmental factors. Inasmuch as both intrapsychic needs and environmental stresses can be reasonably expected to vary for different age groups, it can be assumed that suicidal ideation within the groups will reflect these differences.

The present study is an attempt to test this assumption and to investigate further the relationship of age to behavior in suicidal individuals. The analysis of suicide notes (written shortly before individuals take their lives and sometimes as they die) has proved itself an important method in obtaining data regarding ideation and affect of suicidal persons. The necessarily projective quality of such a communication lends itself well to psychodynamic interpretation. Although only about one suicide in three leaves a note, studies by Shneidman and Farberow (1960) and Tuckman, Kleiner, and Lavell (1959) indicated no significant differences between persons who leave notes and those who do not. Tuckman et al. (1960) investigated the validity of notes and reported an agreement with information from an informant in over 90% of the cases. Reliability of suicide note analysis has also been reported to be acceptably high.

¹ Also at Reiss-Davis Child Study Center, Los Angeles, California. Requests for reprints should be sent to Allen R. Darbonne, Central Research Unit—Building 13, Second Floor West, Veterans Administration Center, Wilshire and Sawtelle Boulevards, Los Angeles, California 90073.

METHOD

The suicide note analysis technique is the method used in the present investigation. Two hundred and fifty nine notes written by 156 individuals were analyzed. This number represents the total number of suicide notes written by men and women who committed suicide during a 3-year period and who resided in the 34 moderately advantaged natural communities of Los Angeles County.

The use of socio-cultural-economic area types as study areas is described in detail by Shneidman and Farberow in their 1960 report of a socio-psychological investigation of suicide. Therefore, only a brief description will be given here. Based on the census data of Los Angeles County, the Los Angeles Welfare Planning Council (1958) studied the county and divided it into 100 relatively stable, homogeneous "study areas." Basis for division was average education, income, home ownership, population, and number of people living alone. These 100 areas were redistributed among nine area types on the basis of "urbanization and social rank."

The group of moderately advantaged natural com-

munities, designated Area Type V, was selected for this investigation. Its 34 communities seemed to represent best the average community of Los Angeles. The total Area Type V population was roughly 1,700,000, which represented 31% of the county's population. It is composed of 31% under age 20, 62% aged 20-65, and 7% over age 65. The annual ratio of number of suicides to population in this area was 14 per 100,000, which was identical with the average suicide ratio for the county.

The suicide note writers were divided into four age groups, 20-39, 40-49, 50-59, and 60 and over. The age groups were selected in an attempt to reflect different stress periods in a lifetime.

The technique employed in note analysis is a minor modification of that used previously by Farberow and Shneidman (1960). The method of analysis involved categorizing each note according to a scheme of five general categories: (a) To whom the note was addressed—interpersonal involvements, dyadic relationships, etc.; (b) Reasons for suicide explicitly stated in the note—conscious reason for victim killing himself; (c) Affect indicated or implied—varied emotional expressions; (d) Content

TABLE 1
OUTLINE FOR SUICIDE NOTE ANALYSIS

<p>A. Addressee of suicide note</p> <ol style="list-style-type: none"> 1. Personal affiliation (family or friend) 2. Impersonal affiliation or no addressee (police, to whom it may concern, or other) <p>B. Reasons stated in suicide note</p> <ol style="list-style-type: none"> 11. No reason stated 12. Ill health, illness, physical disability, symptoms, pain 13. Rejection by another, jilted, unloved, not understood, can't live without you, love triangle (other man, woman) 14. Finances, money, bills, debts, job, occupation, unemployment 15. Ennui, tired of life, as a way out, reached end, couldn't go on 16. No point in living, not worth trying 17. Interest in death, other world, hereafter 18. To "join" a (deceased) loved one 19. Isolation, loneliness 20. Confusion, depression, fear, anxiety 21. Being persecuted, hearing voices, losing mind 22. Sex 23. Other (specify) <p>C. Affect indicated in suicide note</p> <ol style="list-style-type: none"> 31. No affect 32. Hostility, criticism, blame, revenge 33. Absolution of other, giving forgiveness (of specific person) 34. Sorrow, seeking forgiveness (from specific persons) 35. Seeking forgiveness from deity 	<p>C. Affect indicated in suicide note (<i>Continued</i>)</p> <ol style="list-style-type: none"> 36. Self-depreciation, self-derogation, self-criticism, guilt, self-blame (fault) 37. Love, idealization, praise, defense 38. Other (specify) <p>D. Specific content other than affect</p> <ol style="list-style-type: none"> 51. No specific content, other than affect 52. Mention of religion, fate, life, world, death (abstraction) 53. Goodbye, farewell 54. Reference to suicidal act, no one responsible 55. Reference to suicide note 56. Instructions re money, business, power of attorney, funeral expenses 57. Instructions re insurance 58. Instructions re material possessions 59. Instructions re children 60. Instructions re own remains 61. Instructions re notification of others 62. Instructions re message to others 63. Other (specify) <p>E. General focus of the suicide note</p> <ol style="list-style-type: none"> 71. Primarily reason for suicide 72. Primarily affect 73. Primarily instructions 74. Primarily abstractions 75. Primarily content other than affect 76. Primarily reflecting own confusion 77. Extremely short, cryptic, enigmatic 78. Other (specify), no content
---	---

other than affect—victim's main concern just before killing himself; and (e) General focus of the note—overall tenor of the last written communication.

Each general category was subdivided into several more specific descriptive scoring categories by which the notes were analyzed. Table 1 lists these scoring categories. Every note was scored according to all five general categories. The four age groups were then compared with respect to the characteristics of each category.

RESULTS

After each note was assigned a score according to the five general categories listed in Table 1, a chi-square analysis was done to determine the relationship between the four age groups and each of the five general categories. Where a specific item had a frequency so small as to be misleading, it was combined with the "other" category before computation of the degrees of freedom and the chi-square value. The result of this analysis was that "reasons for suicide" and "general focus of the suicide note" were found to be related to age at the .05 level of significance. Table 2 indicates the

TABLE 2

RELATIVE FREQUENCIES OF CATEGORIES WITHIN
SUICIDE NOTES OF DIFFERENT AGE GROUPS

Category	Age group			
	20-39	40-49	50-59	60 and over
Reasons stated				
None			High	Low
Health	Low			High
Rejection or love triangle	High			Low
Tired of life as a way out		High		
Isolation or loneliness				High
Other			High	
Affect indicated:				
None			High	
Hostility			High	
Absolution of others			Low	
Sorrow				High
Seeking deity's forgiveness	Low	High		
Self-depreciation	High			Low
Affection	High			
Other		Low		High
General focus of note:				
Primarily reason				High
Primarily affect	High		Low	
Primarily instructions	Low		High	
Other		Low	High	

specific scoring items (of the statistically significant categories) which were found in a relatively high frequency or a relatively low frequency in one age group as compared to all other age groups.

DISCUSSION

A comparison within each age group of the notes suggests that the categories which are statistically related to age have a certain degree of consistency in meaning. They permit interpretation which may have further implications for suicide prevention and therapeutic approach. In some of the statements which follow, the affect category is also mentioned in relationship to the categories of specific reason and general focus of the suicide note. The affect category is considered in this discussion because there was a very strong trend toward establishment of a relationship between age and affect in the suicide note ($p < .10$). Because of the exploratory nature of the note-analysis technique and the critical nature of the problem under study, it was felt that some reference should be made of this observed trend.

The problems which seem to overwhelm the young adults (20-39) more frequently than the older individuals were related to feelings of rejection and especially problems concerning heterosexual love objects and competition for them. There is a notable absence of reasons related to illness and pain. They could probably cope with the world physically but seemed unable to cope with their problems concerning interpersonal relations, and some of the deeper psychic problems related to sexual needs. At their dying moments these people seemed both willing and able to express affect in their suicide notes. A closer look at this affect indicates it is primarily the internalization of blame and guilt and self-depreciation. At the same time they express love and praise for others. Seeking forgiveness from the deity was less frequent in their notes than in the notes of older suicides. Suicide prevention with this type of individual might be best accomplished through dynamic psychotherapy. This could help the individual gain insight into feelings which stem from early object relationships and which precipitate his recurrent interpersonal problems.

Healthy expression of affect should be encouraged. Where the stress is marital and where it is apparent that the patient loves and praises his spouse but feels unworthy and misunderstood the partners might be brought into the therapy. This is a typical note from this sample, written by a 28-year-old woman:

Dear Mom:

I've tried very hard to hide it, but I'm still as much in love with Bill as always. I don't know how he feels about me, but my life is in his hands now. He knows I can't go on without him and if you read this it means he didn't care enough to stop me. I'm sorry to have failed you this way, but I love Bill so much, I'm not quite sane, let alone intelligent without him. It's better for the babies that I just check out this way than for me to keep on being so unhappy around them. There's a letter somewhere in my grey suitcase for Tim when he is 16.

All my love forever,
Mary

Individuals from ages 40 to 49 who committed suicide left a communication of the vanquished. They seemed unequal to the demands of life and were tired or bored and wanted a way out. They entered the period of the change of life already stating they had reached the end and couldn't go on. This group looked beyond life to the deity and petitioned for forgiveness. A combination of individual and group psychotherapy would possibly be most effective in these cases. The aim here might be a more accurate assessment of the self. It is likely that many of these people have spent their earlier years attempting to achieve above their potential or operating below it. The result is a defeatist feeling and/or guilty thoughts of a wasted life. In either case the future is viewed with pessimism. Closing the gap between self-ideal and self would be an immediate goal of the therapy. The special interest of this age group in the deity should be explored. In some cases, and especially in crisis therapy, important psychological support may be gained from the religious convictions of these people. An example of a note from this age group was written by a 43-year-old man.

Please God forgive me, never again after this is over will I hurt anyone. I love my wife and children too much to continue on not knowing when I'll do something again to hurt them . . . My back will

soon give me real trouble and then I'll be worthless to everyone—Something is wrong with me inside and I've tried to find out just what that something is as far back as I can remember. I'm lost. If I continue as I have in the past my family will not have a chance. Now they can.

I'll pay for all when my maker sees me. I cannot continue to hurt people here on earth and I will not. I'm not weak. I must die.

IIII

P.S. God try to forgive me. I believe in God and always have—a person no good like myself should leave this earth.

With the 50–59-year-old suicides, an interesting finding was presented. Even though responding to a need to write a final communication, they less frequently than any other group gave a reason for their suicide. This group's use of the note to give instructions and to convey information and unemotional messages may indicate a plausible explanation. Expression of affect was found significantly less than in other notes. Possibly these people have arrived at the conclusion that others would not care or understand their feelings. They may have decided it is better not to permit themselves emotional involvement. Another reason could be that these are the individuals who are beginning to experience failing health and other "more rational" reasons for suicide. They have given their act a great deal of thought and within their personal logic have concluded death for themselves. Until more is known about this group, they will present a difficult problem for prevention. Depth therapy seems less likely to succeed here than with the previous groups. These individuals would probably be reluctant to accept involvement in a therapeutic relationship. When the therapist does treat such patients, he may find active environmental manipulation much more his role than usual. If one relates significant world events to the lifetime of these suicides, it is not too difficult to imagine from whence they learned to be realists. These were the Americans who met World War I as teenagers, the Depression during the most productive years of their lives, World War II as they approached middle age, and then saw their children called for the Korean conflict. Certainly treatment would need to consider the effect of such events on the individual. An example of a

note from this age group was written by a 50-year-old woman:

To Whom It May Concern,

I, Mary Smith, being of sound mind, do this day, make my last will as follows—I bequeath my rings, Diamond and Black Opal to my daughter-in-law, Doris Jones and any other of my personal belongings she might wish. What money I have in my savings account and my checking account goes to my dear father, as he won't have me to help him.

To my husband, Ed Smith, I leave my furniture and car.

I would like to be buried as close to the grave of John Jones as possible.

Mary Jane Smith

The age group 60 and over attempted to answer the question "why" more than any other age group. Their reasons more frequently than those of younger suicides cited illness, pain, and physical disability which accompanied their increasing years. They also most often told of loneliness and isolation. Their frequently expressed affect of sorrow and seeking forgiveness from specific persons seemed most often related to having caused trouble to others. There is an absence of statements of self-depreciation and guilt. They apparently retained good thoughts of themselves, but sensed rejection by others. Though these individuals had lived the longest, their notes indicated that they were no more tired of life than any other age groups. This seems to suggest that for individuals at this age, once the situation becomes critical, environmental and milieu therapy combined with medical pain relief would be most useful. Family therapy may also be advantageous. This age group presents a special problem for prevention, if only for the reason that they contribute more than their share to suicide statistics. Preventative programs should take into consideration the psychological side effects of compulsory retirement rules, the importance of educating adults to understand and meet the needs of aging parents, and the need for general medical staffs to be cognizant of psychological aspects of physical ailments among the old and the emotional counterparts of physical aging. This is the note of a 71-year-old man in this group:

John, I do not seem to be wanted living with you and Betty so the best thing I can do is put myself out of the way. When you brought up about the large telephone bill you had sure upset me. I have not been using your phone so I do not understand why you said that . . . So goodbye to you all and please forgive me. I will never get any better at my age so what is the use of trying to go on . . . just a lot of suffering and expense. Say goodbye to Martha Smith for me and treat her right. You know we all have our faults when we get older. She has been very nice to me for what she has done since Mom died. I have had this heart trouble since 1948 and it sure is not going to get better. I do not want to spend several months or years in a rest home. Good Bye. Dad

Two categories of reasons for suicide were relatively absent from the notes of all age groups. These were reasons related to finances and reasons indicating confusion or bizarre psychotic thoughts. Further investigation is necessary to determine the full meaning of this finding.

The evidence presented supports the hypothesis that type of suicidal ideation, including conscious reason for the act, varies with age of the suicidal person. These results are in accord with the view that suicidal behavior is a symptom of various kinds of personality disorganization and not the result of a single suicidal personality syndrome. Attention must be paid to the effect of age variables upon such disorganization in order to enhance scientific understanding and increase effective prevention of suicide.

REFERENCES

- FARBEROW, N. L., & SHNEIDMAN, E. S. Suicide and age. In E. S. Shneidman & N. L. Farberow (Eds.), *Clues to suicide*. New York: McGraw-Hill, 1957.
- SHNEIDMAN, E. S., & FARBEROW, N. L. A socio-psychological investigation of suicide. In H. P. David & J. C. Brengelmann (Eds.), *Perspectives in personality research*. New York: Springer, 1960.
- TUCKMAN, J., KLEINER, R. J., & LAVELL, M. Emotional contents of suicide notes. *American Journal of Psychiatry*, 1959, 116, 59-63.
- TUCKMAN, J., KLEINER, R. J., & LAVELL, M. Credibility of suicide notes. *American Journal of Psychiatry*, 1960, 116, 1104-1106.
- WELFARE PLANNING COUNCIL. *Guides for planning of leisure time services*. Los Angeles: Welfare Planning Council, 1958.

(Received February 26, 1968)

AN OBJECTIVE INSTRUMENT FOR MEASURING DEFENSE MECHANISMS

GOLDINE C. GLESER¹

AND

DAVID IHILEVICH

University of Cincinnati College of Medicine

*Genesee County Community Mental Health
Services Board, Flint, Michigan*

The Defense Mechanism Inventory, a paper and pencil test which purports to measure the relative intensity of usage of five major groups of defenses, is described. The inventory consists of 10 brief stories, two per conflict area, followed by four questions regarding S's actual behavior, fantasy behavior, thoughts, and feelings in the situations described. Five responses typifying the five defenses (i.e., hostility-out, projection, principalization, turning-against-self, and reversal) are provided for each question, from which S selects the one most representative and the one least representative of his reaction. Reliability, validity, and normative studies are presented, all indicating that this inventory has many potential uses as a clinical and research instrument.

Increased preoccupation with the theory of ego processes during the last few decades has led to greater concentration on operational definition, classification, and measurement of various ego functions. Recently, several investigators (e.g., Haan, 1963; Kroeber, 1963) have made progress in differentiating and assessing defensive and adaptive ego functioning. However, there is still much need for objective studies in this area. This preliminary report on the construction and validation of a new instrument designed specifically for the measurement of five general defense mechanisms should be, therefore, of considerable interest to researchers and clinicians.

Many defense mechanisms have been identified, some with rather tenuous differences. This multiplicity has led to attempts to achieve a more parsimonious classification system (e.g., Blum, 1953; Hilgard, 1949; Miller, 1953; Miller & Swanson, 1960). Most such systems, however, have not provided clear-cut criteria by which the various defense mechanisms might be grouped. The need for such a grouping is apparent both in the area of measurement and research and in the clinical situation where one may wish to assess

the major defense mechanisms of an individual. When defenses are evaluated clinically, either from interview material (Raines & Roher, 1955) or from projective test protocols (Filer, 1952), consensus is minimal. In those instances where substantial agreement has been obtained from projective techniques (e.g., Gardiner, Holzman, Klein, Linton, & Spence, 1959), judges inferred a specific set of defenses on the basis of careful training in the scoring procedure.

Only a few objective scales have been developed to measure defenses. The most popular are the Rosenzweig Picture Frustration Test, the Blacky Defense Preference Inquiry, and Byrne's Repression-Sensitization scale. All three have demonstrated some stability over time, but their usefulness as measures of the extent to which certain defenses are employed is still open to question. Fry (1949) and Vane (1954), among others, question the validity of Rosenzweig's (1950) Picture Frustration Test for assessing defenses. With regard to the Blacky Defense Preference Inquiry, Blum (1956) reports that avoidance is the only defense for which some validity has been established. While Byrne's (Byrne, Barry, & Nelson 1963) Repression-Sensitization scale is supported by a number of positive validity studies, a method that provides information on only two defenses is likely to be of limited value either for research purposes or for the clinician. Furthermore, we have found that the Repression-Sensitization

¹The authors wish to thank Huston O. Cagle, director of Adult Psychiatric Clinic, Dayton, Ohio, and Raymond Bodwin, director of Flint Mental Health Clinic, Flint, Michigan, for providing the facilities for collecting part of the data presented in this study. Requests for reprints should be sent to Goldine C. Gleaser, Department of Psychiatry, Cincinnati General Hospital, Cincinnati, Ohio 45229.

scale and Welsh's (1956) Factor A correlate .97 for males and .96 for females in two independent samples of psychiatric outpatients. This finding raises a question as to whether the Repression-Sensitization scale measures any construct other than anxiety.

THE CLASSIFICATION SYSTEM OF THE DMI

Underlying the formulation of the Defense Mechanism Inventory (DMI) is the general assumption that the major function of defenses is the resolution of conflicts between what is perceived by the individual and his internalized values (Kroeber, 1963; Miller & Swanson, 1960). The conflict is resolved by a process whereby the ego attacks, distorts, or becomes selectively unaware of certain aspects of the internal or external world.

On the basis of this a classification system was devised which appears economical and yet is general enough to encompass the most important defense mechanisms previously identified. Five clusters of defenses were delineated.

1. *Turning against Object (TAO)*. This class of defenses deals with conflict through attacking a real or presumed external frustrating object. Such classical defenses as identification-with-the-aggressor and displacement can be placed in this category.

2. *Projection (PRO)*. Included here are defenses which justify the expression of aggression toward an external object through first attributing to it, without unequivocal evidence, negative intent, or characteristics.

3. *Principalization (PRN)*. This class of defenses deals with conflict through invoking a general principle that "splits off" affect from content and represses the former. Defenses such as intellectualization, isolation, and rationalization fall into this category.

4. *Turning against Self (TAS)*. In this class are those defenses that handle conflict through directing aggressive behavior toward S himself. Masochism and autosadism are examples of defensive solutions in this category.

5. *Reversal (REV)*. This class includes defenses that deal with conflict by responding in a positive or neutral fashion to a frustrating object which might be expected to evoke a negative reaction. Defenses such as

negation, denial, reaction formation, and repression are subsumed under this category.

The usefulness of this method of classification has been enhanced by recent findings from the field of perception. Specifically, Witkin and his colleagues report (Witkin, 1965; Witkin, Lewis, & Weil, 1966) that field independent Ss are high on hostility out, projection, or isolation, whereas field dependents Ss are high on either repression or hostility turned inwards. Measurement of these five defenses by means of the DMI should permit further exploration of the relationship between defenses and cognitive styles.

DESCRIPTION OF THE DMI

Twelve stories were developed, two for each of six conflict areas. The conflict areas tapped are authority, independence, masculinity (male form only), femininity (female form only), competition, and situational. The stories were structured in this manner to insure a broad coverage of the areas in which most persons encounter conflicts and to make possible an examination of the notion that persons will differ in the defenses they use according to the nature of the conflict.

The Ss are given either the masculine or the feminine form of the DMI. Each version contains 10 stories. After reading each story S is asked to respond to four questions corresponding to four types of behavior evoked by the situation described in the story: (a) proposed actual behavior, (b) impulsive behavior (in fantasy), (c) thoughts, and (d) feelings. Five responses are provided for each question, each response representing one of the five defense mechanisms listed above. The S marks a plus for the response most representative of his reaction and a minus for that least representative. This structured inquiry makes it possible to examine substantive questions regarding the consistency of a person's defenses over various levels (i.e., actual behavior, fantasy behavior, thoughts, and feelings).

The responses marked with a plus sign by S are given the numerical value of two, those marked with a minus sign are scored zero, and the unmarked responses are given the value of one. Thus the score for any one defense can range from zero to 80, but the

sum of scores for the five defenses must equal 200. A scoring template is superimposed over the answer sheet making it possible to summarize the numerical value of the five defenses across the four levels of behavior in 3-5 minutes. To complete the test itself requires between 30 and 40 minutes.

Preliminary Studies

Five clinical psychologists² were given the names of the six conflict areas and were asked to classify the stories, assigning two to each area. Four accomplished the task in perfect agreement and in accord with the authors' intent. The fifth psychologist did not note that he was to assign two stories to each area and combined the stories involving "authority" and "independence" into one category. All of the others were correctly classified.

The response alternatives have been revised several times. In one study the five previously mentioned psychologists were asked to match each set of five responses with the five defense mechanisms according to the definitions provided. Their judgments agreed with the key for all but 18 responses out of the 240. An examination of these alternatives revealed ambiguities and the responses were duly revised. In another study, the test was administered to 42 Ss and the responses to each question were analyzed to reveal alternatives which were either too popular or unpopular. After this revision, a final check of the inter-item consistency and the popularity (social desirability) of the response alternatives was made using additional samples of 66 male and 52 female students. Equating the alternatives for social desirability tends to minimize the general effect of a set to make a good impression.

When these *DMI* response revisions were completed, another group of clinicians³ was asked to match the responses with the five defense mechanisms in accordance with the definitions provided. One clinician matched all but two responses with that given in the

scoring key; the second, all but four; and the third, all but 10 responses. At least two of the three clinicians agreed with the scoring key on each response.

The effect of free- versus forced-choice instructions on the spread and level of defense mechanism scores and on their generalizability to a universe of content was next investigated. The test was administered to two new samples of male and female undergraduate students at the University of Dayton and the University of Cincinnati. One group was given the forced-choice instructions; the other group was instructed to indicate true or false in front of each response and to put a plus sign in front of the one response that was most representative of the way the respondent would react.

Figure 1 shows the average scores for males and females under free- and forced-choice instructions. The profiles are strikingly similar despite the difference in raw score level. Generalizability of defense scores over a universe of tests structured like this one was ascertained for each sample separately, using a mixed model analysis of variance to take into account the stratified nature of the test (Rajaratnam, Cronbach, & Gleser, 1965). Since the generalizability of scores obtained from a multiple-scored test using free-choice instructions is augmented by any general response tendency, as for example, a tendency to accept or reject relatively few alternatives (Bell, 1962), the coefficients of generalizability obtained under these instructions were corrected by removing the between-persons variance for the average score over all defenses. This correction resulted in more nearly comparable coefficients for the two types of instructions. These coefficients are presented in Table 1. The generalizability of the separate scores are very similar from one sample to another with two exceptions—the *PRN* scale in the male sample using free-choice instructions and the *TAS* scale in the female sample with forced-choice instructions. These coefficients are attenuated by the small variance among persons in the samples tested. The error variances for these scales were no larger than for the remaining scales; hence the low coefficients are probably attributable to sampling fluctuations. Since there was no evi-

² The authors want to express their thanks to William Seeman, George Kisker, Virginia T. Graham, Kayla Springer, and Kenneth Kuntz.

³ The authors want to express their thanks to Michael Hirt, Richard Kurtz, and Leon I. Maizlish.

Males

Females

TABLE 1

GENERALIZABILITY OF *DMI* SCORES DERIVED
FROM INTERNAL CONSISTENCY ESTIMATES

Defense	Males		Females	
	Free response ^a	Forced choice	Free response ^a	Forced choice
<i>TAO</i>	.78	.76	.60	.83
<i>PRO</i>	.78	.57	.75	.63
<i>PRN</i>	-.31	.73	.70	.77
<i>TAS</i>	.63	.71	.73	.21
<i>REV</i>	.77	.59	.56	.65
Total profile	.70	.69	.69	.71

Note.—For Males, Free response, *N* = 15, Forced choice, *N* = 19; for Females, Free response, *N* = 16, Forced choice, *N* = 27.

^a Coefficients are corrected for elevation factor.

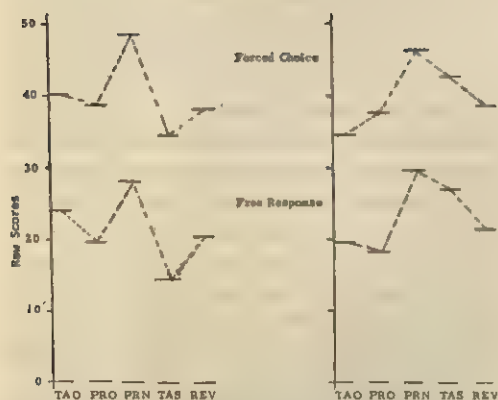


FIG. 1. Mean scores on defenses by sex and type of response.

dence of any loss in differential information of the five defense mechanisms resulting from use of the forced-choice format and it had the advantage of eliminating any response bias not attributable to defenses per se, it was decided to use forced-choice in all further studies.

Distribution of Scores

The means and standard deviations of scores on the five defenses obtained on several

samples of *Ss* are presented in Table 2. Michigan University college students tended to obtain higher scores on *TAO* and *TAS* and lower scores on *PRN* and *REV* than did the general adult *Ss*. These latter were unsystematically selected *Ss* mainly of middle socioeconomic status and employed as probation officers, social workers, vocational counselors, teachers, secretaries, college students, and housewives. Since the adult *Ss* were somewhat older than the college students it is possible that the observed trends are at least partially a function of age. Further evidence that *TAO*

TABLE 2
MEANS AND STANDARD DEVIATIONS OF THE FIVE DEFENSE MECHANISMS
OF THE *DMI* FOR SEVERAL SAMPLES OF *Ss*

Defense	Sophomore college students ^a				General adult subjects				Psychiatric outpatients ^b			
	Males		Females		Males		Females		Males		Females	
	\bar{X}	<i>SD</i>	\bar{X}	<i>SD</i>	\bar{X}	<i>SD</i>	\bar{X}	<i>SD</i>	\bar{X}	<i>SD</i>	\bar{X}	<i>SD</i>
<i>TAO</i>	40.9	8.8	38.9	8.0	39.4	7.8	34.8	8.1	36.8	10.7	33.8	10.4
<i>PRO</i>	39.6	5.9	36.9	6.3	38.4	6.7	36.9	5.4	36.2	6.1	35.4	7.2
<i>PRN</i>	45.3	6.2	45.5	5.9	48.4	6.8	47.3	6.4	45.6	6.2	45.5	7.3
<i>TAS</i>	37.5	6.9	44.0	6.9	34.4	7.6	41.9	4.9	39.6	7.4	44.2	7.9
<i>REV</i>	36.6	7.2	35.1	7.6	39.6	6.3	39.2	6.8	41.9	9.4	41.1	10.7
Age					23.0	9.2	28.7	9.2	32.6	8.3	31.6	7.8
Education					13.8	2.1	13.4	1.8	12.3	2.3	12.1	1.9

Note.—For Sophomore college students, males, *N* = 226, females, *N* = 180; for General adult *Ss*, males, *N* = 43, females, *N* = 71; for Psychiatric outpatients, males, *N* = 124, females, *N* = 110.

^a University of Michigan.

^b Combined data for all outpatients to whom *DMI* was administered.

decreases with age while *PRN* and *REV* increase is indicated by the correlations between age and each of these scores in male and female outpatient samples. (See Table 3.)

In both the college and the general adult samples and also in the psychiatric outpatient samples, males are significantly higher than females on *TAO* and significantly lower on *TAS*. Males are also consistently higher on *PRO* than females, and significantly so in the student sample.

The male psychiatric patients tend to have lower scores on *TAO* and *PRO* and higher scores on *TAS* and *REV* than do either of the nonpatient samples.

Stability of Defenses

The stability of defense scores over time was investigated in two small samples of male and female Ss. One group consisted of 12 counselors who were administered the *DMI* before and after a T-group experience, an interval of a week. The product-moment correlations between the separate defense scores ranged from .85 for *PRO* to .93 for *TAO*. The average correlation was .89. The second sample consisted of 11 first-year psychology graduate students who were administered the *DMI* at the beginning and end of a course with 3 months intervening. The correlations for this sample ranged from .69 for *PRN* to .87 for *TAO* with an average of .76. These results are encouraging, but further studies of stability are needed.

CONSTRUCT VALIDATION

In order to provide a more stringent validation of our classification system we provided a list of 15 defenses⁴ in alphabetical order to three psychologists and seven social workers and asked them to match each of the 240 responses of the *DMI* with one defense from the list. The judges were also allowed to mark in front of a response the name of a defense not mentioned in the list or note that in their opinion the response did not represent a defense mechanism. Analysis of the data indi-

cated that there was satisfactory agreement with the key (more than 60%) on those responses keyed *TAS*, *REV*, and *PRN* but not on responses keyed *TAO* and *PRO*. Almost a third of the judgments regarding *TAO* responses were to the effect that these responses did not constitute a defense, while 19% of the judgments of responses keyed *PRO* indicated that they were not defenses and another 14% were identified as defenses which we subsume under *TAO*. These results indicate that further revision is needed on some of the responses in these two defense categories.

Relationship among Defenses

It was expected that the five defense groupings would interrelate systematically both in accordance with predictions derived from psychoanalytic theory (Bellak, 1958; Fenichel, 1945; Hartmann, 1952; Rapaport, 1951) and with previously published findings. From an empirical standpoint it was hoped, however, that these correlations would not be so high as to imply that only a single factor was being measured.

The intercorrelations obtained for the previously described samples of college students, unselected normal Ss, and outpatients, each stratified by sex, are presented in Table 3. The matrices are remarkably similar, indicating considerable stability in the relationship of defenses from sample to sample.

TAO and *PRO* are positively correlated to a moderate degree (.29 to .63). This relationship is in agreement with the findings of Caine (1960) and Peak, Muney, and Clay (1960). Lesser (1958) reports a positive correlation between hostility and anti-Semitism, presumably a form of projection. From a theoretical standpoint a positive relationship was expected since both modes of defense entail the expression of aggression. The only other positive correlation among the five defenses is between *PRN* and *REV* (.33 to .68). There is a paucity of pertinent research findings here. A related finding, however, is that of Chodoff, Friedman, and Hamburg (1964), who reported that parents of leukemia patients tended to employ denial and isolation of affect as their major defenses. On a theoretical basis, repression of affect is common to

⁴ Avoidance, denial, displacement, identification with the aggressor, intellectualization, isolation, negation, projection, rationalization, reaction formation, repression, regression, reversal, turning against self, undoing.

TABLE 3
INTERCORRELATIONS AMONG THE FIVE DEFENSES OF THE *DMI*
FOR THREE SAMPLES STRATIFIED BY SEX

Sample	TAO	PRO	PRN	TAS	REV		
I. University of Michigan college students							
TAO		.46	-.51	-.39	-.64		
PRO	.45		-.48	-.37	-.58		
PRN	-.60	-.34		-.24	.40		
TAS	-.41	-.42	-.21		-.07		
REV	-.67	-.66	.45	.06			
II. General adult population							
TAO		.29	-.58	-.23	-.69		
PRO	.31		-.44	-.17	-.56		
PRN	-.66	-.61		-.29	.33		
TAS	-.31	-.15	-.56		-.09		
REV	-.71	-.55	.68	-.03			
III. Outpatient							
	TAO	PRO	PRN	TAS	REV	Age	Education
TAO		.53	-.63	-.26	-.70	-.27**	-.02
PRO	.63		-.55	-.26	-.62	-.06	-.03
PRN	-.60	-.48		-.28	.54	.21	.05
TAS	-.44	-.45	-.14		-.19	-.08	-.06
REV	-.79	-.64	.47	.00		.23*	.05
Age	-.39**	-.18	.24	.11	.33*		-.00
Education	.28*	.16	.00	-.22	-.22	-.05	

Note.—For the University of Michigan sample, $N = 111$ females upper triangle, $N = 226$ males lower triangle; for General adult population, $N = 71$ females upper triangle, $N = 43$ males lower triangle; for Outpatient sample, $N = 93$ females upper triangle, $N = 67$ males lower triangle.

* $p < .05$.
** $p < .01$.

both types of defenses, which would lead one to expect a positive relationship.

Both *TAO* and *PRO* are substantially negatively correlated with *PRN* and with *REV* ($-.44$ to $-.79$). Negative correlations of hostility or projection with some of the specific defenses subsumed under reversal have been reported by Goldstein (1952), Peak et al. (1960), and Shipman and Marquette (1963). The negative correlations between hostility or projection and principalization have not been reported elsewhere to our knowledge but might be predicted on the basis that both hostility and projection are considered more primitive, immature responses to conflict than is principalization (Bellak, 1958; Fenichel, 1945; Rapaport, 1951). *TAS* is negatively correlated to some slight degree with each of the other defenses except *REV*. *TAS* and *REV* are independent. Relevant research on

the relationship of *TAS* to other defenses is meager and even conflicting. For example, Mussen and Naylor (1954) found that aggression directed outward and that directed inward tend to coexist to a large degree, while Purcell (1956) reports findings of a negative relationship between them.

Relationship between Scores on the DMI and MMPI Scales

MMPI data were obtained on 67 male and 93 female psychiatric outpatients who had also been administered the *DMI*. The MMPI responses were scored on all original scales as well as for some of the better-known experimental scales. Complete matrices of intercorrelations were obtained for males and females separately. A portion of these results is shown in Table 4.

All of the *DMI* scales except *PRO* have a

sizable number of MMPI correlates. *TAO* is positively correlated with scores on *F*, *Pd*, *Sc*, and *Ma* and negatively correlated with *L*. Furthermore, *TAO* is positively correlated with *Mf* and *A* for males but not for females. *PRO* has a somewhat similar pattern of correlations with the MMPI scales but the correlations are smaller and generally not significant for one or the other or both sexes. *PRN* and *REV*, on the other hand, are negatively correlated with *F*, *Pd*, *Pa*, *Pt*, and *Sc*. Social introversion (*Si*) and anxiety (*A*) are negatively correlated with *REV* for males and with *PRN* for both sexes. Both these scales are positively correlated with *TAS* as is *D*. Scores on Barron's (1953) ego-strength scale are negatively correlated with *TAS*. These patterns of correlations are for the most part consistent with theoretical predictions.

The correlates of *L* and *K* are of particular interest since these scales are sometimes considered to be measures of a set on the part of *S* to appear socially desirable. In general, *L* and *K* are negatively correlated with *TAO* and *PRO*, and positively correlated with *REV*, particularly in the male sample.

If the correlations are attributed to a common factor of social desirability it would imply that the more socially desirable responses in the *DMI* are the *REV* alternatives while the least desirable are the *TAO* and *PRO* responses. But then we have the peculiar situation that psychiatric outpatients obtain a more socially desirable profile than do general adults or college sophomores. (See Table 2.) A more likely explanation for the correlations is that they are due to a common factor of evasiveness or denial (versus forthrightness) which is revealed both in *S*'s attitude about himself (MMPI) and his handling of conflict situations (*DMI*).

Relationships between Defenses from the DMI and Haan's Defense Scales

Haan (1965) developed scales for eight defenses using MMPI items. Her definitions differed somewhat from ours and the scales have never been cross-validated but it seemed reasonable to explore their relationship to the *DMI* scales. Hence scores were obtained on these scales for the above-mentioned psychiatric clinic patients and also for a sample of

TABLE 4
CORRELATIONS BETWEEN *DMI* AND MMPI SCALES FOR 67 MALE
AND 93 FEMALE PSYCHIATRIC OUTPATIENTS

MMPI variables	<i>DMI</i> Scales									
	<i>TAO</i>		<i>PRO</i>		<i>PRN</i>		<i>TAS</i>		<i>REV</i>	
	M	F	M	F	M	F	M	F	M	F
<i>L</i>	-.40**	-.31**	-.37**	-.12	.16	.19	.05	.02	.55**	.24*
<i>F</i>	.25*	.22*	.12	.24*	-.14	-.42**	.10	.37**	-.36**	-.36**
<i>K</i>	-.21	-.01	-.29*	-.17	.18	.30**	-.11	-.26**	.39**	.09
<i>Hs</i>	-.17	-.04	-.20	-.12	.04	-.09	.07	.24*	.23	-.01
<i>D</i>	-.03	-.04	-.10	.02	-.17	-.13	.42**	.24*	-.15	-.06
<i>Hy</i>	-.03	-.04	-.06	-.08	.02	-.06	.06	.16	.01	.01
<i>Pd</i>	.19	.23*	.02	.12	-.11	-.26*	.08	.18	-.23	-.23*
<i>Mf</i>	.38**	.09	.24*	-.08	-.08	.14	-.05	-.14	-.50**	-.05
<i>Pa</i>	.19	.01	.07	.08	-.10	-.26*	.06	.46**	-.25*	-.21*
<i>Pt</i>	.18	.07	-.02	.00	-.20	-.33**	.22	.36**	-.25*	-.12
<i>Sc</i>	.34**	.18	.05	.14	-.27*	-.41**	.10	.41**	-.34**	-.28**
<i>Ma</i>	.20	.37**	.08	.16	.06	-.31**	-.23	-.01	-.21	-.11
<i>Si</i>	.16	-.05	.06	.04	-.24*	-.22*	.32**	.32**	-.35**	-.04
<i>A</i> (Welsh)	.25*	-.03	.11	.10	-.25*	-.31**	.23	.36**	-.39**	-.06
<i>Es</i> (Barron)	.20	.13	.13	.04	.01	.14	-.34**	-.35**	-.03	-.02

Note.—Decimal points omitted for ease of reading.

* $p < .05$.

** $p < .01$.

23 male and 24 females patients at another outpatient clinic. Those correlations which are consistent across samples and significant in at least one sample are the following: Haan's Denial scale is positively correlated with *REV* and *PRN* and negatively correlated with *TAO*. Doubt is positively correlated with *TAS* and negatively with *PRN*. Repression is positively correlated with *REV* and negatively correlated with *TAO*. Primitive Defense is positively correlated with *PRN* and *REV*. One possibly interesting sex difference is that *TAS* and Repression are positively correlated for males and negatively correlated for females. (See Table 5.)

Other Validity Studies

Responses of male alcoholics to the DMI. Michigan State University graduate students conducted a study of the *DMI* responses of 54 male alcoholics (Aldridge, Baxter, Nopziger, Roggenbuck, Shimansky, & Wolthuis, 1967). Relying on psychoanalytic theory to the effect that the alcoholic personality is rooted in a self-destructive, orally dependent orientation to life (Fenichel, 1945; Zwerling & Rosenbaum, 1959) and on the empirical findings that alcoholics demonstrate low ego strength (Borowitz, 1964), self-destructiveness (White, 1966), and dependency (Witkin, Karp, & Goodenough, 1959), the investigators predicted that the alcoholics would be high

on *TAS* and *REV* and low on *TAO* and *PRN* relative to the normals. These predictions were all substantiated at an alpha level of .05 with the exception of *PRN*, for which $p < .10$. The alcoholics also deviated more frequently from the intermediate range (2 *SD* above or below the mean of standardization group) on at least one defense mechanism relative to the normal sample ($p < .001$).

The relationship of DMI to dream recall. The 180 University of Michigan female students on whom mean scores are reported in Table 3 were administered a questionnaire related to dream recall and were asked to keep a dream diary for 1 month, noting in it every morning upon awakening various aspects of their dreams.⁸ Using the dream data drawn from the questionnaire, two predictions were substantiated. The *Ss* ($N = 29$) who reported frequent dream recall (3-4 times per week or more) were higher on *TAO* ($p < .05$) and lower on *REV* ($p < .05$) than were low dream recallers (i.e., eight *Ss* who reported they dream at most a few times per year). These findings are in line with those reported by Tart (1962) and others who attribute infrequent dream recall to the use of repression against threatening dream content. Significant correlations were not obtained, however, between defenses and frequency of dream recall when data from the dream diary were used. One possible interpretation of these results is that people who report low dream recall on a questionnaire are doing so as a result of a continuous process of repression rather than as a consequence of immediate repression following the dream process.

Psychotherapy and DMI scores. In an exploratory study on the effect of psychotherapy on defense mechanisms, we asked each of five experienced psychotherapists to test two or three of his patients, a man and woman preferably, who had been in treatment with him for approximately 6 months. We compared these results with those obtained from a group of patients, matched for education, sex, age, and race (Caucasian) drawn from the waiting list. The results revealed (see Table 6) that patients on the

TABLE 5

CORRELATIONS BETWEEN SCORES ON THE DEFENSE SCALES OF THE *DMI* AND THOSE OF HAAN

Scales	Males		Females	
	Clinic 1	Clinic 2	Clinic 1	Clinic 2
<i>TAO</i> × Displacement	.12	.04	.00	-.05
<i>TAO</i> × Denial	.00	-.40**	-.43*	-.26*
<i>TAO</i> × Repression	-.14	-.33**	-.33	-.20
<i>PRO</i> × Projection	.06	-.04	.36	.01
<i>PRN</i> × Intellectualizing	.29	.09	-.21	.19
<i>PRN</i> × Doubt	-.50**	-.17	-.28	-.26*
<i>PRN</i> × Denial	.34	.33**	.26	.48**
<i>PRN</i> × Primitive Defense	.22	.33**	.24	.07
<i>TAS</i> × Doubt	.33	.26*	.35	.30*
<i>TAS</i> × Repression	.29	.35**	-.10	-.36**
<i>REV</i> × Doubt	-.01	-.25*	-.08	-.10
<i>REV</i> × Denial	.24	.48**	.50**	.30**
<i>REV</i> × Primitive Defense	.23	.37**	.48**	-.08

Note.—For Males, Clinic 1, $N = 23$, Clinic 2, $N = 67$; for Females, Clinic 1, $N = 24$, Clinic 2, $N = 93$.

* $p \leq .05$.

** $p \leq .01$.

⁸ The authors wish to thank David Cohen for permitting the use of this data.

waiting list were significantly high on *REV* (male $p < .02$; female $p < .01$) and low on *PRO* (male $p < .05$; female $p < .01$) relative to patients in treatment. The difference on *REV* might indicate a reduction of denial, repression, and other avoidance mechanisms for patients in therapy, probably as the result of the therapists stressing and encouraging their patients to confront their feelings and face up to their problems. The other finding, that the patients in treatment were higher on *PRO*, was unexpected. One possible explanation is that the group of patients that stayed in treatment for 6 months was a more pathognomonic group relative to the waiting list patients (Grummon as quoted in Seeman, 1965), many of whom may drop out before treatment begins or within a few sessions.

The relationship of DMI to field articulation. Witkin and his colleagues (Witkin, Dyk, Faterston, Goodenough, & Karp, 1962), rely-

ing on what they call the "differentiation hypothesis," propose a network of association between degree of field articulation and various other measures of differentiation (e.g., body concept, sense of identity, cognitive styles, and defenses). Such defenses as denial, repression, and hostility-turned-inward reflect, according to Witkin, a lower degree of differentiation than do isolation, projection, and hostility-turned-outward. Lesser differentiation is assumed in the first group of defenses because their operation involves a more primitive self-structure, as well as lesser separation of self from nonself. In a recently completed study, Ihilevich (1968) related the *DMI* defenses to the cognitive style of field articulation. A sample of 110 psychiatric outpatients (50 males and 60 females) were administered the *DMI*, the Embedded Figure Test, and the Figure Drawing Test. As predicted from Witkin's "differentiation hypothesis," Ss who relied mainly on "global" defenses (*TAS* and *REV*) were more field dependent relative to Ss who relied excessively on "differentiated" defenses (*TAO* and *PRO*) ($p < .01$).

TABLE 6

MEANS AND STANDARD DEVIATIONS OF THE FIVE DEFENSE MECHANISMS OF THE *DMI* FOR A GROUP OF PSYCHIATRIC OUTPATIENTS IN TREATMENT FOR 6 MONTHS AS COMPARED TO A GROUP OF PATIENTS ON THE WAITING LIST

Scale	In therapy		On waiting list		p
	\bar{X}	SD	\bar{X}	SD	
Males					
Age	28.7	6.64	26.0	6.72	
Education	12.8	1.40	13.2	2.18	
<i>TAO</i>	44.2	9.88	38.8	7.65	ns
<i>PRO</i>	42.3	7.09	35.7	5.98	$< .05$
<i>PRN</i>	41.3	5.66	44.3	5.57	ns
<i>TAS</i>	40.7	10.06	40.5	5.50	ns
<i>REV</i>	31.6	8.07	40.6	7.57	$< .02$
Females					
Age	29.4	6.65	26.1	4.85	
Education	12.4	2.17	11.7	2.41	
<i>TAO</i>	39.3	11.34	30.8	11.02	$< .10$
<i>PRO</i>	39.3	7.44	32.5	3.47	$< .01$
<i>PRN</i>	44.4	7.87	48.0	9.79	ns
<i>TAS</i>	42.9	7.94	46.2	6.32	ns
<i>REV</i>	34.1	9.29	42.5	3.70	$< .01$

Note.—Patients on waiting list had been interviewed, diagnosed, and accepted for treatment. For Males in Therapy, $N = 12$, on Waiting List, $N = 11$; for Females in Therapy, $N = 14$, on Waiting List, $N = 11$.

REFERENCES

- ALDRIDGE, R. G., BAXTER, G. I., NOPFIZER, L. M., ROGGENBUCK, A. L., SHIMANSKY, R. L., & WOLTHUIS, D. J. Defense mechanisms of an alcoholic population as compared to a normal population. Unpublished master's thesis, Michigan State University, 1967.
- BARRON, F. An ego strength scale which predicts response to psychotherapy. *Journal of Consulting Psychology*, 1953, 17, 327-333.
- BELL, R. Q. Isolation of elevation and scatter components in personality and attitude questionnaires. *Educational and Psychological Measurement*, 1962, 22, 699-713.
- BELLAK, L. The schizophrenic syndrome. In L. Bellak (Ed.), *Schizophrenia: A review of the syndrome*. New York: Logos Press, 1958.
- BLUM, G. S. *Psychoanalytic theories of personality*. New York: McGraw-Hill, 1953.
- BLUM, G. S. Defense preferences in four countries. *Journal of Projective Techniques*, 1956, 20, 33-41.
- BOROWITZ, G. H. Some ego aspects of alcoholism. *British Journal of Medical Psychology*, 1964, 37, 257-263.
- BYRNE, D., BARRY, J., & NELSON, D. The revised repression-sensitization scale and its relationship to measures of self-description. *Psychological Reports*, 1963, 13, 323-334.
- CAINE, T. M. The expression of hostility and guilt in melancholic and paranoid women. *Journal of Consulting Psychology*, 1960, 24, 18-22.

- CHODOFF, P., FRIEDMAN, S. B., & HAMBURG, D. A. Stress, defenses and coping behavior: Observations in parents of children with malignant disease. *American Journal of Psychiatry*, 1964, 120, 743-749.
- FENICHEL, O. *The psychoanalytic theory of neurosis*. New York: Norton, 1945.
- FILER, R. M. The clinician's personality and his case reports. *American Psychologist*, 1952, 7, 336. (Abstract)
- FREY, F. D. A study of reactions to frustration in 236 college students and in 207 inmates of state prisons. *Journal of Psychology*, 1949, 34, 27-30.
- GARDINER, R. W., HOLZMAN, P. S., KLEIN, G. S., LINTON, H., & SPENCE, D. P. Cognitive control: A study of individual consistencies in cognitive behavior. *Psychological Issues*, 1959, 1, No. 4.
- GOLDSTEIN, S. A projective study of psychoanalytic mechanisms of defense. Unpublished doctoral dissertation, University of Michigan, 1952.
- HAAN, N. Proposed model of ego functioning: Coping and defense mechanisms in relationship to I.Q. change. *Psychological Monographs*, 1963, 77(8, Whole No. 571).
- HAAN, N. Coping and defense mechanisms related to personality inventories. *Journal of Consulting Psychology*, 1965, 29, 373-378.
- HARTMANN, H. Mutual influences in the development of ego and id. *Psychoanalytic study of the child*, 1952, 7, 42-50.
- HILGARD, E. R. Human motives and the concept of self. *American Psychologist*, 1949, 4, 374-382.
- IHILEVICH, D. The relationship between defense mechanisms and field dependence-independence. Unpublished doctoral dissertation, University of Cincinnati, 1968.
- KROEBER, T. The coping functions of the ego mechanisms. In R. White (Ed.), *The study of lives*. New York: Atherton Press, 1963.
- LESSER, G. S. Extrapunitiveness and ethnic attitude. *Journal of Abnormal and Social Psychology*, 1958, 56, 281.
- MILLER, D. R. *Lectures in genetic psychology*. University of Michigan, 1953.
- MILLER, D. R., & SWANSON, G. R. *Inner conflict and defense*. New York: Holt, 1960.
- MUSSEN, P. H., & NAYLOR, H. K. The relationship between overt and fantasy aggression. *Journal of Abnormal and Social Psychology*, 1954, 49, 235-240.
- PEAK, H., MUNKE, B., & CLAY, M. Opposite structures, defenses, and attitudes. *Psychological Monographs*, 1960, 74(8, Whole No. 495).
- PURCELL, K. The thematic apperception test and anti-social behavior. *Journal of Consulting Psychology*, 1956, 20, 449-456.
- RAINES, G. N., & ROHRER, E. H. The operational matrix of psychiatric practice. I. Consistency and variability in interview impressions of different psychiatrists. *American Journal of Psychiatry*, 1955, 114, 721-733.
- RAJARATNAM, N., CRONBACH, L. J., & GLESER, G. C. Generalizability of stratified-parallel tests. *Psychometrika*, 1965, 30, 39-56.
- RAPAPORT, D. Toward a theory of thinking. In D. Rapaport (Ed.), *Organization and pathology of thought*. New York: Columbia University Press, 1951.
- ROSENZWEIG, S. Revised norms for the adult form of the Rosenzweig P-F Study. *Journal of Personality*, 1950, 18, 344-346.
- SEEMAN, J. Perspectives in client-centered therapy. In B. B. Wolman (Ed.), *Handbook of clinical psychology*. New York: McGraw-Hill, 1965.
- SHIPMAN, W. G., & MARQUETTE, C. H. The manifest hostility scale: A validation study. *Journal of Clinical Psychology*, 1963, 19, 104-105.
- TART, C. T. Frequency of dream recall and some personality measures. *Journal of Consulting Psychology*, 1962, 26, 467-470.
- VANE, J. R. Implications of the performance of delinquent girls on the Rosenzweig P-F Study. *Journal of Consulting Psychology*, 1954, 18, 414.
- WELSH, G. S. Factor dimensions A and R. In G. S. Welsh and W. G. Dahlstrom (Eds.), *Basic readings on the MMPI in psychology and medicine*. Minneapolis: University of Minnesota Press, 1956.
- WHITE, W. F. Personality and cognitive learning among alcoholics with different intervals of sobriety. *Psychological Reports*, 1965, 16, 1125-1140.
- WITKIN, H. A. Psychological differentiation and forms of pathology. *Journal of Abnormal Psychology*, 1965, 70, 317-336.
- WITKIN, H. A., DYK, R. B., FATERSON, F. H., GOODENOUGH, D. R., & KARP, S. A. *Psychological differentiation: Studies of development*. New York: Wiley, 1962.
- WITKIN, H. A., KARP, S. A., & GOODENOUGH, D. R. Dependence in alcoholics. *Quarterly Journal of Studies in Alcoholism*, 1959, 20, 493-504.
- WITKIN, H. A., LEWIS, H. B., & WEIL, E. Shame and guilt reactions of more differentiated and less differentiated patients early in therapy. In S. S. Tomkins (Chm.), *Some recent work on shame and guilt*. Symposium presented at the meeting of the American Psychological Association, New York City, September 1966.
- ZWERLING, I., & ROSENBAUM, M. Alcoholic addiction and personality. In S. Arieti (Ed.), *American handbook of psychiatry*. Vol. I. New York: Basic Books, 1959.

(Received February 28, 1968)

EGO FUNCTIONS IN DISTURBED AND NORMAL CHILDREN: ASPIRATION, INHIBITION, TIME ESTIMATION, AND DELAYED GRATIFICATION¹

ANTHONY DAVIDS²

Brown University

A battery of procedures for assessing ego functions was administered to normal elementary school children and institutionalized emotionally disturbed children, subdivided into 3 age levels. Orderly, predictable developmental differences in ego functioning were found within the normal sample, but there was little regularity or consistency within the abnormal sample. Comparison of the normal and abnormal groups revealed no differences in motor manipulation or level of aspiration, but the disturbed children were less able to inhibit motor responsivity, less accurate in time estimation, and less willing to seek delayed gratification in preference to immediate reward. The findings are in keeping with expectations based on psychoanalytically oriented theories of ego development. Their implications for early diagnosis, prevention, and education were discussed.

In recent years, an increasing amount of psychological research has been devoted to studies of ego functions. Singer (1955) presented a comprehensive review of theories of ego development and indicated the implications of these theories for clinical and experimental research. Singer and his colleagues (Singer, Wilensky, & McCraven, 1956) also reported empirical findings from studies of basic ego functions such as delaying capacity, fantasy, and planning ability in adult schizophrenics. Levine, Spivack, Fuschillo, and Tavernier (1959) and Spivack, Levine, and Sprigle (1959) published reports of research on ego functioning in which they focused on intelligence, motor inhibition, time sense, and self-control in disturbed adolescents. Similar phenomena have also been studied by Siegman (1961) in a group of young prisoners. Cross-cultural studies of delay of gratification in normal children have been conducted by

Mischel (1959, 1961). Davids and his colleagues (Davids, Kidder, & Reich, 1962; Davids & Parenti, 1958; Davids & Sidman, 1962) studied time orientation in emotionally disturbed children and institutionalized juvenile delinquents. Recently, Davids and Sidman (1962) reported findings from a study of impulsivity, time orientation, and delayed gratification in underachieving and high achieving secondary school students.

Most of these independent investigations show either an explicit or an implicit link with psychoanalytic theory (Freud, 1920, 1946) and they provide empirical evidence pertinent to Rapaport's thesis (1950, 1951) that development of thought processes in childhood results from inevitable delay of gratification and the necessity of learning to inhibit impulse expression. On the basis of these psychoanalytically oriented theories of ego development and previous empirical findings, the following predictions were formulated for examination in the present study: (a) basic ego functions are less well developed in younger children than in older children, and (b) at each developmental level, emotionally disturbed children evidence less adequate ego functioning than do normal children.

Although the present investigation is related conceptually to several of the studies cited above and employs experimental pro-

¹ This investigation was facilitated by small grant M-4327 (A) from the National Institute of Mental Health. The author wishes to acknowledge the assistance of Robert B. Carlin and Alexander W. Carnathan who collected the data for this study. The author is indebted, also, to Lewis P. Lipsitt who arranged for the testing of normal children from the Seekonk, Massachusetts, public schools. The emotionally disturbed children were studied at the Emma Pendleton Bradley Hospital in Riverside, Rhode Island.

² Also at the Emma Pendleton Bradley Hospital, Riverside, Rhode Island.

cedures similar to those utilized independently in some of these varied researches, it is designed to provide findings that have not been available previously. The experimental design is more comprehensive in that it includes both normal and disturbed Ss at younger developmental levels than those studied previously and includes a wider variety of procedures for assessing ego functions. More specifically, the present study includes normal children from public schools and children who were institutionalized for emotional disturbances. Within both samples, the children were subdivided into three developmental levels and were administered a battery of experimental procedures designed to assess ego functions such as goal setting, motor performance, control of impulsivity, time estimation, and delay of gratification.

METHOD

Subjects

This study employed two samples of children—34 normal boys who were attending public school and 38 emotionally disturbed boys who were undergoing residential psychiatric treatment. The normal sample was subdivided into the following three groups: thirteen 7-year-olds in the second grade, ten 9-year-olds in the fourth grade, and eleven 11-year-olds in the sixth grade. The average IQ of the children in each of these grade levels was between 100 and 110, with an overall mean of 105. The emotionally disturbed sample was subdivided into the following three groups: four 7-year-olds, thirteen 9-year-olds, and twenty-one 11-year-olds. In the disturbed sample, the average IQ for each age group was between 90 and 100, with an overall mean of 96. Many of these disturbed children do not perform on intelligence tests at a level in keeping with their estimated potential based on clinical evaluation (Davids, 1958). Therefore, the sample of disturbed children was also subgrouped on the basis of their mental ages as follows: eleven 7-year-olds, thirteen 9-year-olds, and fourteen 11-year-olds. The mean mental age in both the overall group of normal children and the overall group of disturbed children was identical (10.0 years). In analyzing findings obtained in this research, separate comparisons were made for subgroupings based on chronological age and mental age. Since essentially identical results were obtained from the two methods, the results presented in this report will be based on chronological age subgroupings within the disturbed and normal samples. It should be emphasized, however, that the small number of younger children in the abnormal group may limit the generality of findings obtained.

The attempt was made to include all testable boys

within the residential treatment setting. This group of disturbed boys included four schizophrenic children and four who were classified as neurotic, but the vast majority ($N=30$) were classified as cases of severe behavior disorder (passive-aggressive personality). These are the kind of boys who are termed "acting-out" children—hostile, antisocial, explosive, and often unable to learn in the conventional school situation.

Measures

The children were seen individually, and the procedures were administered in the order described below.

Rate of manipulation and level of aspiration. The child was presented with the Minnesota Rate of Manipulation Form Board and was instructed to turn over as many of the blocks as he could before being told to stop. Following this 30-second trial, *E* recorded the number turned and announced it to the child. He then asked the child to state how many he thought he could turn over if he were given another trial. The level of aspiration was recorded, but the child was not permitted a second trial, as the present experimental design attempted to eliminate the effect of success or failure on the remainder of the experimental procedures. Since some of the children were unable or unwilling to state their level of aspiration in terms of a specific number, but gave responses such as "I'll do more next time," or "I'll do the same as (or less than) I just did," the responses were dichotomized into the categories of "better" and "same or worse." The data from this procedure were analyzed in terms of these two global measures of aspiration.

Motor inhibition. The child was presented with a mimeographed sheet containing a spiral. He was told to place his pencil at the center of the spiral and to draw his way out (in a manner similar to tracing a path through a maze), going as slowly as he possibly could, but without stopping the movement of the pencil. In other words, there was a winding path from the center of the spiral to the outside, and the child was supposed to traverse the path as slowly as possible but with continuous motor movement. The time to complete this task was recorded, and a maximum allowable time was set at 2 minutes.

Time estimation. A stopwatch was shown to the child and he was told that his task was to estimate a 30-second passage of time. The *E* then started the stopwatch, hidden from the child's view, and stopped the watch when the child indicated what he believed to be 30 seconds. The actual elapsed time was recorded and announced to the child, and he was then given a second trial, in order to measure improvement or decrement.

Delay of gratification. The child was asked the following questions. First, "If I gave you 10¢, what would you do with it?" Then, "If I gave you \$1, what would you do with it?" The answers were recorded verbatim and, for each of the two amounts, were later classified into the following three cate-

gories: (a) immediate gratification (e.g., buy candy, buy a coke, etc.), (b) short delay of gratification (e.g., go to movies on the weekend, go to a skating rink within a few days, etc.), (c) longer delay of gratification (e.g., put it in the bank, save it toward a bicycle or a dog, etc.).

RESULTS

Rate of Manipulation and Level of Aspiration

There is a consistent developmental progression indicating increased efficiency in motor manipulation. In the normal group, the three means (youngest = 20, middle = 30, oldest = 37) are significantly different. In the abnormal group, the youngest and middle subgroups do not differ significantly (23 vs. 26), but both of them perform significantly less efficiently than the oldest subgroup ($M = 35$). In comparing the normal and abnormal groups at each of the three developmental levels, no significant differences were obtained. The mean number of blocks turned in the overall abnormal group was 30, while the mean for the normal group was 28, indicating that psychopathology had no marked effect on this measure of motor performance.

The level of aspiration findings reveal no developmental consistency. In the normal group, the youngest subgroup contained the largest percentage of Ss (54%) stating they would do "better" on a second trial. In the abnormal group, the oldest subgroup contained the largest percentage of Ss (57%) aspiring to improve on a second trial. Within both the normal and abnormal groups, none of the differences between percentages (ranging from 40 to 57%) were statistically significant. Comparing the total normal group with the total abnormal group reveals a non-significant difference between the proportions (53% vs. 47%) aspiring to do "better" on the second attempt at motor manipulation. Thus, in the present experimental situation the aspirations of emotionally disturbed and normal children were very similar.

Motor Inhibition

The findings presented in Table 1 reveal several noteworthy differences between the normal and abnormal groups. Within the normal sample, there is an orderly develop-

mental trend indicating a decrease in impulsivity and increase in motor control with increasing age. The two younger subgroups do not differ significantly, but both of these subgroups obtain significantly lower scores than the oldest subgroup. Within the abnormal sample, there is no orderly developmental progression. In fact, the youngest subgroup in this sample obtained the highest mean score on the motor inhibition task. None of the differences between means in this abnormal sample, however, are statistically significant. Comparisons between the normal and abnormal samples reveal highly significant differences at both the middle and oldest age levels. Comparison of the scores in the two overall samples reveals that the mean score in the normal sample was almost twice as high as the mean in the abnormal sample. These findings, then, provide definite evidence of greater impulsivity and less motor control in the emotionally disturbed children.

Time Estimation

The results obtained from the two attempts at time estimation reveal several similarities and some significant differences between the disturbed and normal groups of children. On the first trial, within the normal group, the youngest and the middle age groups make significantly higher error scores than do the children in the oldest subgroup. Within the abnormal group, there is a trend indicative of decreasing error scores on the first time-estimation trial as a function of increasing developmental levels. However, there are no significant differences among the error scores in the abnormal sample. As shown in Table 1, at each of the developmental levels, the disturbed children receive higher error scores than do the normal children, but the only statistically significant difference is found at the oldest age level. At this developmental level, the mean error score for the disturbed children is 13.6, which is markedly higher than the mean error score of 7.7 in the normal children.⁸

⁸ These mean error scores are in the direction of underestimation. That is, they indicate the average amount of time by which the children responded before the 30 seconds had elapsed. In both the dis-

TABLE 1

COMPARISON OF SCORES (IN SECONDS) FOR NORMAL AND ABNORMAL GROUPS ON
MOTOR INHIBITION AND TIME ESTIMATION

Item	Normal group		Abnormal group		<i>F</i>	<i>t</i>
	<i>M</i>	Variance	<i>M</i>	Variance		
Motor inhibition						
Youngest	43.3	520.2	40.0	2846.0	5.47*	.12
Middle	46.6	393.4	24.2	499.7	1.27	2.50*
Oldest	76.8	2079.9	29.3	1086.1	1.92	3.21**
Combined	55.1	1160.6	28.3	1000.9	1.16	3.44**
Time-estimation error						
Trial 1						
Youngest	13.5	94.4	16.3	90.9	1.04	.50
Middle	13.7	69.1	14.2	102.4	1.48	.13
Oldest	7.7	40.4	13.6	55.3	1.37	2.23*
Combined	11.6	73.1	14.1	71.1	1.03	1.22
Trial 2						
Youngest	13.8	60.7	15.5	123.0	2.03	.34
Middle	11.2	92.4	14.5	86.8	1.06	.84
Oldest	4.2	18.8	9.4	72.1	3.84*	2.32*
Combined	9.8	70.3	11.8	84.4	1.20	.96

Note.—For the normal group, *N* = 34; for the abnormal group, *N* = 38.**p* < .05.***p* < .01.

On the second time-estimation trial, again there are significant differences within the normal sample, with the two younger subgroups obtaining significantly higher error scores than the oldest subgroup. On this second trial, within the disturbed sample there is also a trend indicative of decreasing error scores with increasing age, but none of the differences within this sample are significant. Comparison of the disturbed and normal children, as shown in Table 1, reveals higher error scores for the disturbed children at each of the three developmental levels, although the difference is statistically significant only at the oldest age level in this study. At this developmental level, the mean error score in the disturbed subgroup is over twice as high as the mean error score in the normal children.

Comparison of each child's error scores (in seconds) on the two trials provides a measure

turbed and normal groups, very few children erred in the direction of waiting beyond 30 seconds before responding. Thus, although errors could be made in either direction, almost every child in this study responded too soon.

of improvement on the basis of this one-trial learning. That is, if a child obtains a lower error score on the second trial it is assumed that he learned something from his performance on the first trial and that this one-trial learning leads to improvement on the second trial. Within both the disturbed and normal samples there is an orderly progression indicating increased improvement scores as a function of increasing developmental levels. However, none of the differences among the mean improvement scores (.7, 2.7, 4.5) are significant within the normal sample and the only statistically significant difference within the abnormal sample (.8, 1.2, 4.2) is between the middle and oldest subgroups. Comparison of the disturbed and normal children at each developmental level reveals no statistically significant differences between the groups. In general, then, the findings indicate rather consistent tendencies for older children to be more accurate judges of the passage of brief intervals of time and for normal children to be more accurate than emotionally disturbed children. Although few of the differences within and between the groups are highly sig-

nificant statistically, there is consistency and orderliness within the findings, and in many instances the actual magnitude of the differences is noteworthy.

Delay of Gratification

As shown in Table 2, the normal group contained a significantly greater proportion who would delay gratification if they were given 10¢. Whereas 50% of these normal children would spend the money immediately, 76% of the disturbed children would seek immediate gratification. At each of the three age levels, there were more disturbed children who would spend the 10¢ immediately, although the difference between proportions is significant only in the comparison between the youngest subgroup in the disturbed and normal samples. Chi-square analyses within the disturbed group and within the normal group revealed no significant association between the age groupings and the three categories of delay of gratification. In fact, in the normal sample, the youngest subgroup contained the smallest proportion of children who would seek immediate gratification, while in the disturbed sample the youngest subgroup contained the highest percentage of children who would *not* delay gratification as measured by this procedure. As mentioned previously, however, findings from the youngest group of abnormal children must be viewed with caution, and may well be of lim-

ited generality, due to the small number of Ss in this category in the present study.

Analysis of findings utilizing the \$1 value, instead of the 10¢ value, again reveal a higher proportion of disturbed children seeking immediate gratification at each of the three age levels. As shown in Table 2, for the total sample, 26% of the normal children would spend \$1 immediately and 50% of the disturbed children would spend this amount immediately. Just as with the 10¢ value, chi-square analysis of within-group relations between age levels and the three categories of gratification reveal no significant associations either within the normal sample or within the disturbed sample.

It is evident from these findings that both groups contain a greater proportion of children who would delay gratification when the amount is \$1 than when the amount is 10¢. The differences between the proportions showing immediate gratification with these two amounts is statistically significant in both samples. For the normal group, the difference between the proportions spending 10¢ and spending \$1 is significant at the .04 level and for the disturbed group the difference is significant at the .02 level.

Thus it can be concluded that there is no consistent developmental trend within either the disturbed group or the normal group, for either the 10¢ or the \$1 value, but that for each of these two values a greater proportion

TABLE 2
COMPARISON OF NORMAL AND ABNORMAL GROUPS ON DELAY OF GRATIFICATION

Amount	% normal group			% abnormal group			Critical ratio
	Spend	Save short	Save long	Spend	Save short	Save long	
Small (10¢)							
Youngest	38	8	54	100	0	0	2.18*
Middle	60	0	40	69	0	31	.45
Oldest	55	9	36	76	10	14	1.20
Combined	50	6	44	76	5	19	2.30*
Larger (\$1)							
Youngest	31	23	46	50	50	0	.69
Middle	10	30	60	54	8	38	2.20*
Oldest	36	18	46	48	14	38	.65
Combined	26	24	50	50	16	34	2.10*

* $p < .05$.

of the disturbed children would seek immediate gratification and more of the normals would save the money for some future gratification. Also, for both emotionally disturbed and normal children there is a realistic perception of the difference in value between the amounts of 10¢ and \$1, and increasing the amount of money involved has the effect of reducing the proportion of children who would seek immediate gratification in both the disturbed and normal samples.

Associations among Measures

As shown in Table 3, the only significant correlate of IQ in the group of normal children is the time-estimation error score on Trial 1. Motor manipulation is associated significantly with both motor inhibition and errors on the second time-estimation trial, indicating that those normal children who turned the most blocks tended to spend longer time drawing their way out of the spiral and were more accurate in their time estimations. For this normal group, the correlations between time-estimation and motor inhibition scores seem particularly noteworthy. The correlation of .88 between the two trials of the time-estimation task indicates a high degree of reliability. Moreover, there are significant negative associations between time-estimation scores, on both trials, and the measure of motor inhibition. These coefficients indicate that, within the normal group, those children

who were most accurate in their judgments of the passage of time were better able to delay their motor responses in a situation calling for control and inhibition. It is evident from Table 3 that, for the normal Ss, the level of aspiration and delay of gratification measures are not significantly associated with the other measures. However, there is highly significant association between the two measures of delay of gratification, with those normal children who would save 10¢ tending to be those who would also save \$1.

As shown in Table 4, in the abnormal group, IQ is positively associated with motor manipulation, level of aspiration, and time-estimation improvement, and is negatively associated with time-estimation errors on the second trial. In other words, those disturbed children who did better on the IQ test, turned more blocks, evidenced a higher level of aspiration, were more accurate in estimating the passage of time, and showed greater improvement from one trial to the next. Motor manipulation is negatively associated with time-estimation errors and positively associated with time-estimation improvement score. Thus, those disturbed children who obtained higher scores in the block-turning situation tended to be more accurate in the time-estimation situation, and also showed greatest improvement from the first to the second trial.

There is a correlation of .78 between the two measures of time estimation, indicating a

TABLE 3
ASSOCIATIONS AMONG MEASURES WITHIN THE NORMAL GROUP

Measure	2	3	4	5	6	7	8	9
IQ (1)	.19	.23	-.43*	-.31	-.32	.12	.00	1.14
Motor manipulation (2)		.35*	-.14	-.35*	.17	.00	.12	.52
Motor inhibition (3)			-.50**	-.42*	-.07	.50	.12	.95
Time-estimation error 1 (4)				.88**	.27	2.95	.00	2.72
Time-estimation error 2 (5)					-.21	.12	.00	2.64
Time-estimation improvement (6)								
Level of aspiration (7)						.07	.00	2.71
Delay of gratification (10¢) (8)							.00	1.91
Delay of gratification (\$1) (9)								9.54**

Note.—The values shown under Columns 2 through 6 are product-moment coefficients of correlation; those shown under Columns 7 and 8 are chi-square tests of association, with $df = 1$, corrected for continuity; those shown under Column 9 are chi-

* $p < .05$.

** $p < .01$.

TABLE 4
ASSOCIATIONS AMONG MEASURES WITHIN THE ABNORMAL GROUP

Measure	2	3	4	5	6	7	8	9
IQ (1)	.64**	.08	-.28	-.42**	.38*	3.89*	.02	.67
Motor manipulation (2)		-.06	-.29	-.46**	.43**	1.22	.34	4.56
Motor inhibition (3)			-.02	-.27	.38*	1.22	.03	1.96
Time-estimation error 1 (4)				.78**	.08	.00	.90	.45
Time-estimation error 2 (5)					-.52**	.10	.58	.13
Time-estimation improvement (6)						.00	.03	.69
Level of aspiration (7)							.03	3.09
Delay of gratification (10¢) (8)								3.70
Delay of gratification (\$1) (9)								

Note.—The values shown under Columns 2 through 6 are product-moment coefficients of correlation; those shown under Columns 7 and 8 are chi-square tests of association, with $df = 1$, corrected for continuity; those shown under Column 9 are chi-squares with $df = 2$. $N = 38$.

* $p < .05$.

** $p < .01$.

satisfactory degree of reliability. This finding is in keeping with the relationship found in the normal group. However, whereas significant correlations were found between the time-estimation error scores and the motor inhibition scores for the normal children, there is no significant association between time errors and motor inhibition in the abnormal group. Interestingly, however, there is a significant correlation between time-estimation improvement score and motor inhibition score in the abnormal group. That is, those disturbed children who showed the greatest improvement in their time estimations from one trial to the next tended to be the ones who were better able to comply with the instructions calling for motor inhibition and control. As with the normal group, there is little evidence of associations between the level of aspiration and delay of gratification measures and the other measures within the abnormal group of children.

It may be that the simple goal-setting task employed in the present study is not sufficiently refined to show consistently high relations with various other measures designed to assess ego functions. Moreover, it should not be overlooked that the delay of gratification procedures employed in this study were restricted to two values, with the children being asked to respond to a hypothetical situation. If the procedure employed different values to be either gained immediately or

postponed to the future and/or if the children were confronted with actual choices, rather than hypothetical choices, then it might be that significant associations would be found among measures of preference for delayed gratification and other signs of ego strength. These possibilities remain to be explored in future studies.

DISCUSSION

The findings from this investigation reveal noteworthy similarities and differences in varied aspects of ego functioning in normal children and emotionally disturbed youngsters. In the area of motor manipulation these two groups of children do not differ and, within both samples, there is orderly progression of better performance as a function of age. Previous research (Davids & Oliver, 1960; Davids & White, 1958) has also demonstrated that on simple tasks requiring gross motor movements emotionally disturbed children perform as well as, or even better than, normal children.

Interestingly, however, when the task calls for a voluntary slowing down of response, with inhibition of motor movements, the disturbed children function much less effectively than do the normal children. In other words, when the situation demands that they work rapidly and with abandon, the disturbed children are successful, but when they are asked to restrain their energies and to inhibit their motor

responses they fare less well than do normal children. It is most impressive that a task as simple as drawing a path around a spiral differentiates to such an extent between children who function adequately in ordinary society and children whose behavior in everyday life is so socially unacceptable that they have to live in the special environment offered by a residential psychiatric treatment center.

Since these disturbed children have so much difficulty in complying with the demands of this very simple uncomplicated task, it is obvious that they would experience great difficulty in the usual classroom situation. Many of the children in the disturbed sample would be characterized as suffering from the hyperkinetic impulse disorder (Laufer, Denhoff, & Solomons, 1957) and most of them are at least 2 years retarded in scholastic attainment (Talmadge, Davids, & Laufer, 1963). They have low frustration tolerance and are extremely restless and active in an ordinary class situation. This sort of behavior makes it difficult to cope with the demands of the usual learning situation, and as these children proceed through the early developmental years they fall further behind in conventional learning and mastery of academic material. The present experiment demonstrates quite convincingly that motor inhibition is hard for them to master and this area of difficulty must be considered in any educational planning for children like these. If one cannot even get them to slow down or to inhibit their responses in such a simple situation, how can they be expected to cope with complex phenomena like reading and writing as they are taught conventionally in the public schools?

Although the level of aspiration situation constituted only a small facet of this experiment and the goal-setting task employed may seem overly simple, one might well be surprised to find that the disturbed children do not differ from the normal children. In an earlier study, using different groups of disturbed and normal children, Davids and White (1958) found that under neutral conditions disturbed children set goals for themselves that were at least as high as those established by normal children. However, if they met failure or criticism, their aspiration

levels suffered to a much greater extent than did those evidenced by normal children.

An interesting finding, in this regard, is the significant association between the level of aspiration measure and IQ scores for the disturbed children, with absolutely no relationship between these measures in the normal group. Other studies (Davids, 1958) have shown that emotional disturbance often has a pronounced detrimental effect on intellectual functioning in children. It seems that many of the abilities called for on conventional intelligence tests are similar to those we have attempted to study in the present research. Within the abnormal group, those children who are highly motivated to turn the blocks in keeping with the examiner's instructions to work rapidly, who desire to do even better on a second performance as evidenced by their stated aspiration, and who have more accurate time orientation, are the ones who also tend to obtain higher scores on the IQ test. With the normal children, however, there is noticeably less association between these measures of ego functions and IQ scores.

The failure of regularity in development of time orientation is also demonstrated by the present study. Within the normal sample there are marked differences between the time-estimation accuracy of younger and older children, but the differences between younger and older emotionally disturbed children are much less striking and are nonsignificant. According to psychoanalytic theory, a prime function of the ego is to integrate events from the past, present, and future of a person's life space. The normal individual must be able to inhibit impulsivity and forsake immediate pleasures in order to attain future gratifications that accompany the fulfillment of long-range goals. Since the disturbed children are believed to possess weak ego structure, they should show impaired temporal orientation in comparison with their age mates who possess more adequate ego structures. This investigation provides empirical evidence that this is, in fact, the case.

The findings in regard to delay of gratification are also in keeping with theoretical expectations, with a greater proportion of normal children, at each of the developmental

levels incorporated in the present experimental design, evidencing the ability to delay gratification. Most theories of child development propose that the ability to delay gratification has its origin in the early feeding situation between mother and child. The child learns to wait for his rewards and comes to learn that food (and love) will be forthcoming. This requires the development of a sense of basic trust in one's environment (Erikson, 1950). In the microcosm of the present study we see that emotionally disturbed children, whose case histories attest to early infantile feeding difficulties and considerable hostility and insecurity in the mother-child relationship (Davids & Lawton, 1961), are much more disposed to seek immediate gratification and are less oriented toward more distant goals.

The problem for the future is how to prevent these ego disturbances from developing and, once developed, how to effect changes in a personally and socially enhancing direction. Studies of ego functions in adults in mental hospitals, prisons, and other institutional settings will undoubtedly shed further light on the differences between disturbed and normal individuals. However, such studies can provide only retrospective evidence for pathological antecedents and, unfortunately, reveal little in the way of effecting basic changes. It is hoped that further studies which improve upon and extend the kind of information gathered from the present experimental approach will lead not only to increased understanding of ego functions in children but also to improved methods for early modification of inadequate ego development before the damage becomes irreversible.

REFERENCES

- DAVIDS, A. Intelligence in childhood schizophrenics, other emotionally disturbed children, and their mothers. *Journal of Consulting Psychology*, 1958, 22, 159-163.
- DAVIDS, A., KIDDER, C., & REICH, M. Time orientation in male and female juvenile delinquents. *Journal of Abnormal and Social Psychology*, 1962, 64, 239-240.
- DAVIDS, A., & LAWTON, M. J. Self concept, mother concept, and food aversions in emotionally disturbed and normal children. *Journal of Abnormal and Social Psychology*, 1961, 62, 309-314.
- DAVIDS, A., & OLIVER, G. R. Learning, performance, and retention under stress in emotionally disturbed and normal children. *Journal of Personality*, 1960, 28, 368-381.
- DAVIDS, A., & PARENTI, A. N. Time orientation and interpersonal relations of emotionally disturbed and normal children. *Journal of Abnormal and Social Psychology*, 1958, 57, 299-305.
- DAVIDS, A., & SIDMAN, J. A pilot study—impulsivity, time orientation, and delayed gratification in future scientists and in underachieving high school students. *Exceptional Children*, 1962, 29, 170-174.
- DAVIDS, A., & WHITE, A. A. Effects of success, failure, and social facilitation on level of aspiration in emotionally disturbed and normal children. *Journal of Personality*, 1958, 26, 77-93.
- ERIKSON, E. H. *Childhood and society*. New York: Norton, 1950.
- FREUD, S. *A general introduction to psychoanalysis*. New York: Boni & Liveright, 1920.
- FREUD, S. Formulations regarding the two principles in mental functioning. In *Collected papers*. Vol. IV. London: Hogarth, 1946.
- LAUFER, M. W., DENHOFF, E., & SOLOMONS, G. Hyperkinetic impulse disorder in children's behavior problems. *Psychosomatic Medicine*, 1957, 19, 39-49.
- LEVINE, M., & SPIVACK, G. Incentive, time conception and self control in a group of emotionally disturbed boys. *Journal of Clinical Psychology*, 1959, 15, 110-113.
- LEVINE, M., SPIVACK, G., FUSCHILLO, J., & TAVERNIER, A. Intelligence and measures of inhibition and time sense. *Journal of Clinical Psychology*, 1959, 15, 224-226.
- LEVINE, M., SPIVACK, G., & WRIGHT, B. The inhibition process, Rorschach human movement, responses, and intelligence: Some further data. *Journal of Consulting Psychology*, 1959, 23, 306-312.
- MISCHEL, W. Preference for delayed reinforcement: An experimental study of a cultural observation. *Journal of Abnormal and Social Psychology*, 1959, 56, 57-61.
- MISCHEL, W. Delay of gratification, need for achievement, and acquiescence in another culture. *Journal of Abnormal and Social Psychology*, 1961, 62, 543-552.
- RAPAPORT, D. On the psychoanalytic theory of thinking. *International Journal of Psychoanalysis*, 1950, 31, 161-170.
- RAPAPORT, D. *Organization and pathology of thought*. New York: Columbia University Press, 1951.
- SIEGMAN, A. W. The relationship between future time perspective, time estimation, and impulse control in a group of young offenders and in a control group. *Journal of Consulting Psychology*, 1961, 25, 470-475.
- SINGER, J. L. Delayed gratification and ego development: Implications for clinical and experimental research. *Journal of Consulting Psychology*, 1955, 19, 259-266.

- SINGER, J. L., WILENSKY, H., & MCCRAVEN, V. G. Delaying capacity, fantasy, and planning ability: A factorial study of some basic ego functions. *Journal of Consulting Psychology*, 1956, 20, 375-383.
- SPIVACK, G., LEVINE, M., & SPRIGLE, H. Intelligence test performance and the delay function of the

ego. *Journal of Consulting Psychology*, 1959, 23, 428-431.

- TALMADGE, M., DAVIDS, A., & LAUFER, M. W. A study of experimental methods for teaching emotionally disturbed, brain damaged, retarded readers. *Journal of Educational Research*, 1963, 56, 311-316.

(Received March 4, 1968)

Manuscripts Accepted for Publication in the

Journal of Consulting and Clinical Psychology

This listing plus those published in preceding issues represents the backlog of manuscripts accepted by this Journal. Such listing is intended to allow readers to become aware of research many months in advance of journal publication.

- Childhood Parental Relationships of Homosexual Men: Ray B. Evans*: Department of Psychiatry, Loma Linda University, Loma Linda, California 92354.
- Childhood Parental Relationships and the Establishment of Gender Roles of Homosexuals: Ralph H. Gundlach*: 160 East 84th Street, Apt. 2K, New York, New York 10028.
- Parental Relations and Male Homosexuality in Patient and Nonpatient Samples: Evelyn Hooker*: Department of Psychology, University of California, Los Angeles, Los Angeles, California 90024.
- Self-Concept of the Married Psychiatric Patient and His Mate's Perception of Him: Martin Harrow*, David A. Fox, and Thomas Detre: Department of Psychiatry, Yale University, 333 Cedar Street, New Haven, Connecticut.
- Relationship of Selected Psychosocial Variables to Prognostic Judgments: James E. Cowden* and Asher R. Pacht: 1 West Wilson Street, P. O. Box 669, Madison, Wisconsin 53701.
- Psychiatric Hospital Experience and Attitudes toward "Mental Illness": J. J. Smith*: Kings View Hospital, P. O. Box 631, Reedley, California 93654.
- Development of an Empathy Scale: Robert Hogan*: Department of Psychology, The Johns Hopkins University, Baltimore, Maryland 21218.
- Personality Characteristics of Therapists: Description of Relevant Variables and Examination of Conscious Preferences: Aron Z. Spilken*, Martin A. Jacobs, James J. Muller, and Jane Knitzer: Boston University School of Medicine, Division of Psychiatry, 80 East Concord Street, Boston, Massachusetts 02118.
- Reality Attentiveness-Inattentiveness and Externalization-Internalization in Defensive Style: Marshall Bush*, Robert Hatcher, and Martin Mayman: Department of Psychiatry, Mount Zion Hospital, 1600 Divisadero Street, San Francisco, California 94115.
- Word Association Repertoires of Schizophrenics and Normals: George D. Fuller and Solis L. Kates*: Department of Psychology, University of Massachusetts, Amherst, Massachusetts 01003.
- Validation Study of Marital Status and the Self-Report Scale for Process-Reactive Schizophrenia: Donald Meichenbaum*: Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada.
- Bodily Concerns and the WISC Object Assembly Subtest: Gerald Marsden* and Neil Kalter: Departments of Psychology and Psychiatry, University Hospital, Ann Arbor, Michigan.
- Situational Appraisal Inventory: Development and Validation of a Measure of Evalu-

*Asterisk indicates author for whom address is supplied.

PRELIMINARY SOCIALIZATION FOR PSYCHOTHERAPY OF BEHAVIOR-DISORDERED ADOLESCENTS

ROBERT L. SCHWITZGEBEL¹

University of California, Los Angeles

A common problem in treating adolescent delinquents is their failure to participate in traditional psychotherapeutic procedures. Small cash bonuses given contingent on participation in tape-recorded nondirective interviews were shown to be effective in shaping dependable and prompt attendance. Using a variable schedule of reinforcement, extrinsic rewards could be gradually reduced without corresponding decrement in attendance behavior. The rationale for using such a procedure is emphasized.

A patient cannot be treated in absentia. A minimum expectation in traditional psychotherapy is that the patient will come to a specified location at a predetermined time on a fairly regular basis for a person-to-person verbal exchange. And generally it is assumed that voluntary participation is highly desirable, if not essential, for success of the treatment (cf. Hollingshead & Redlich, 1958).

Unfortunately, a substantial number of persons whose behavior often represents some social threat (e.g., psychotics, psychopaths, adolescent delinquents) may be among the least likely to volunteer for therapy. In a follow-up study of 54 adolescents with behavior disorders referred to a psychiatric clinic, Hammar and Holterman (1965) found that 21 would-be patients failed to appear, and 11 more discontinued after the initial contact. Rosen, Bahn, Shellow, and Bower (1965) reported that of an estimated 194,000 adolescents treated in outpatient clinics in the United States in 1962, approximately one-third withdrew from treatment. Similar lack of adequate voluntary attendance is reflected in employment data cited by Gordon (1965). Of 234,000 underprivileged youths invited by letter to visit an employment counselor, only 42,000 appeared for interviews. Of these, fewer than 13,000 were referred to jobs, and less than 7,000 were hired—some for only a few days. Similar to orthodox psychotherapists, the typical employer expects or demands a high rate of goal behavior (e.g., "After all,

I'm running a business."). Failure of the youth to meet this criterion usually results in threats and punishment which may antagonize the youth and cause him to devalue himself further as a potential employee.

The statistics which have been cited simply confirm a persistent and widely acknowledged difficulty confronting virtually every practitioner who deals with this population. Anna Freud (1958), for example, wrote that adolescents are very difficult to get into treatment because they do not cooperate, they miss appointments, they are unpunctual, they cannot or will not introspect, and their rapidly changing emotional patterns leave little energy available to invest in the analyst. Stieper and Wiener (1965) have claimed that developing motivation in behavior-disordered patients "is probably the single most crucial problem with which therapy must grapple [p. 123]." The purpose of this paper is to summarize briefly a strategy for dealing with adolescent delinquents and certain other groups of prospective patients who are often minimally motivated to participate in psychotherapy.

RATIONALE

The reasons usually given for poor participation involve the adolescent's desire for independence, revolt against authority, a narcissistic belief in his ability to solve his own problems, reluctance to talk about hostile or sexual thoughts, etc. (e.g., Pelpz, 1957). In Aichhorn's classic work, *Wayward Youth* (1955), he argued that one of the chief difficulties in treatment was the fact that a delinquent "does not suffer discomfort from his symptoms [p. 27]" in the way which a neu-

¹Requests for reprints should be sent to Robert Schwitzgebel, Department of Psychology, University of California, Los Angeles, Los Angeles, California 90024.

rotic does. Alternately, one might hypothesize that traditional psychotherapeutic procedures simply appear more threatening to some potential patients than others. In this respect, it should be noted that many of the characteristics attributed here to delinquents which tend to make them unsuitable for traditional psychotherapy may be class-linked rather than related to age or pathology. That is, these predispositions characterize a large number of persons, especially in the lower socioeconomic class, who are neither adolescent nor delinquent (cf. Riessman, Cohen, & Pearl, 1964).

Therapy procedures which seemingly require adolescent delinquents (particularly males) to acknowledge feelings of weakness, ignorance, helplessness, or passivity may be extremely aversive. From the adolescent's point of view, accepting help may imply that he is weak or sick or "queer." He is likely to believe that psychotherapists are a strange kind of doctors who do mysterious things to crazy people. The delinquent knows that he is not like the crazy people he has seen in television movies. To the contrary, he may pride himself on being active, strong, and clever. This image is especially valued in his peer group subculture. In exchange for an adult's promise of something better in the future, it seems to the delinquent that he must sacrifice, in advance, his social identity and his place on the street corner. These may not be much, but they are all he has.

Should the delinquent happen to volunteer or be forced into an initial interview, the experience may be embarrassing and punitive because traditional mental health procedures are not designed to treat patients with his characteristic reactions. For example, he may, having had no other experience, relate to the psychotherapist as he has to (other) medical doctors. He waits for questions, gives brief answers (not always accurate), and waits for the next question or for some specific advice. The therapist, on the other hand, who is often a trainee or resident in the clinic, may rigidly wait for the patient to verbally elaborate on feelings. The ensuing silences and awkward exchanges are likely to confirm the patient's doubts regarding the possibility of getting "real help," and the therapist may become more convinced that psychotherapy is appro-

priately restricted to individuals who exhibit what Schofield (1964, p. 133) has referred to as the "Yavis" syndrome—persons who are youthful, attractive, verbal, intelligent, and successful.

Unfortunately, the patient's willingness to become involved has often been attributed to factors (e.g., "motivation" or "intelligence") which seem to lie outside the province of treatment. Attendance at psychotherapeutic interviews is not, however, an intrinsic attribute of prospective patients. It is the result of innumerable variables such as physical setting and the nature of verbal and nonverbal communications as well as the patient's predispositions. Cooperation is more practicably viewed not as "something" which the patient *brings to* therapy but as a *product of* therapy.

Since cooperative behavior among delinquents is likely to be minimal and of short duration, the therapist may have only one or two opportunities to strengthen the desired responses. Therefore some relatively simple, direct, but flexible strategies are mandatory. Again, Aichhorn (1955) appropriately noted:

I consider this first moment of our coming together of the utmost importance. It is more than a "feeling out" of the situation; it must have the appearance of certainty and sureness and must be put through as quickly as possible because in most cases it forms the foundation for our later relationship. The adolescent does the same thing when he comes in contact with me. He wants to know right away what kind of person he is dealing with [p. 99].

Passivity on the part of the therapist may be misinterpreted as weakness, lack of concern, or hostile withholding of information. On the other hand, gestures of friendship and attention such as a sincere compliment about clothes or appearance may actually prove threatening. This is particularly true in the case of adolescent delinquents whose personal histories confirm that attention from elders is usually paired with punishment. The occasions on which parents, teachers, or public officials have given attention to our prospective patient *without* some form of punishment may be indeed rare. We can assume that fear and counter-hostility will generalize to the therapy situation unless the therapist's actions are sufficiently unique to force the delinquent to make new discriminations. Thus, one of the

primary tasks of the initial interview is to break up stereotyped response patterns with minimal aversiveness to the patient.

STRATEGIES

Despite theoretical differences, a survey of the techniques actually employed during the initial phases of treatment of behavior-disordered adolescents by a number of well-known therapists shows remarkable conformity. Slavson (1956) noted, for example, that in group work, "elements of surprise and novelty were constantly being introduced into the group therapy situation in the form of variations in activity and unexpected refreshments [p. 20]." Aichhorn (1955) reported that

I usually begin with a friendly look or attitude, sometimes I say, 'How do you do?' or I may only shake hands in silence. I say that there is nothing here to be afraid of, that this is neither a police station nor a court. Sometimes I tell a joke by way of introduction [p. 99].

Schmideberg (1949), who will "make a point of shaking hands with some of these patients as often as possible" and accept an invitation to tea, has claimed that her "spontaneous and unprofessional manner helps to break down their depersonalization which is an important defense mechanism of the 'callous' criminal [p. 186]."

Neill (1962) has provided an example of what might be considered positive reinforcement of attendance behavior at his psychoanalytically oriented interviews known as "private lessons":

"With young children, the technique was more spontaneous. I followed the child's lead. Here is a typical first P.L. with a six-year-old girl named Margaret. She comes into my room and says, 'I want a P.L.' 'Righto,' I say.

She sits down in an easy chair.

'What is a P.L.?' she asks.

'It isn't anything to eat,' I say, 'but somewhere in this pocket I have a caramel. Ah, here it is.' And I give her the sweet.

'Why do you want a P.L.?' I ask.

'Evelyn has one, and I want one too.'

'Good. You begin it. What do you want to talk about?' [p. 42]"

Eissler (1949) wrote, "It is a general rule of therapy of the delinquent never to act the way he expects, but unceasingly to introduce new and unforeseen elements in order to keep

alive his interest in the therapeutic situation [p. 19]." Specifically, he recommends the use of money, although he cautions that the situation "is wrought with danger and therefore requires special skill and tact."

Gradually, as certain commonalities emerge, successful treatment should become less a matter of "special skill" and more a matter of definable technique. As a rather simple demonstration of a way in which dependable attendance at psychotherapeutic sessions could be obtained, a program of instrumental conditioning was instituted among adolescent delinquents by the present writer. No referral procedure was involved. Prospective Ss were contacted randomly in poolhalls, amusement centers, etc. Inasmuch as details of the experimental results have been reported elsewhere (Schwitzgebel, 1967), this report will emphasize initial socialization procedures which have, over a period of years, become sufficiently systematized and successful to eliminate the problem of obtaining an adequate number of nonincarcerated delinquents for research and clinical training purposes.

METHOD AND RESULTS

Specification of Goal Behavior

The terminal or goal behavior in the present demonstration was dependable and prompt arrival at a clinic office for tape-recorded interviews.

Estimation of Goal Behavior Base Rate

The Ss were selected from a population with a low probability of dependable attendance at a clinic. Twenty-one male adolescent offenders, ranging in age from 12 to 21 years and averaging 15 months of probation, were contacted directly on street corners or in amusement centers by Es and asked to participate in tape-recorded interviews. It was explained that S's task would be to express his opinion and feelings about parents, teachers, policemen, schools, reformatories, etc.; they would be paid \$1.00 an hour, but one of the "qualifications" for the job was a court record. (Records were verified later.)

Eight members of the experimental group had histories of active refusal to participate in psychiatric treatment, and seven Ss had been incarcerated an average of 18 months. The base rate for attendance among Ss for the first four interviews showed a group mean of 38 minutes late. (Failure to arrive for the scheduled appointment was scored as 60 minutes; nine out of 21 Ss failed to appear for the first interview, and approximately 50% of those who did arrive were late.) Most Ss who missed an interview were able to be contacted a few days later at the original meeting place.

Selection of Reinforcing Consequences

Eissler (1949) has claimed that "usually no successful therapy of a delinquent is possible without his obtaining money from his analyst during that introductory phase of treatment when all possibilities must be mobilized to make the analyst a valuable person in the delinquent's eyes [p. 20]." Money is not only a readily-exchanged generalized reinforcer for most delinquents, but it frequently serves to "protect" the prospective patient from criticism by the peer group. Almost any activity is approved in the delinquent subculture if the participant can "make a fast buck"—this includes visits to a psychotherapist. Other tangible "face-saving" consequences or valued activities will sometimes suffice. If money is used in the initial phase, it can be gradually replaced by more personalized and idiosyncratic reinforcers. These usually prove to be more effective and are well-received by the patient.

Premack (1963, 1965) has suggested that any activity which has a higher probability of occurring can be used as a reinforcer for a response of lesser probability. Thus, if at a given time, the independent response rates in a "free choice" situation for an adolescent would show the following rank order: Watching a movie (M) > playing pool (P) > eating candy (C) > reading a comic book (B), then P could reinforce C and B but not M. If true, this principle suggests that literally hundreds of potential reinforcers are available for use by the therapist. An obvious difficulty arises, however, in estimating the respective probabilities of various activities unless one has knowledge of previous deprivations. In work with retarded children, Addison and Homme (1966) have used what they term a "behavior menu." The "menu" consists of a booklet of pictures and descriptions of brief activities (e.g., a child being pushed down the hallway on a chair by E) which the child may choose when he becomes eligible for reinforcement. Another simple way of determining what tangible consequences might serve as reinforcers is to take a leisurely walk through a department store with the patient and to note various items which receive favorable comment.

In addition to administrative, ethical, and financial considerations in selecting reinforcers, it is generally held that the reinforcing activity should not be directly incompatible with the goal behavior. For example, dismissing a child from school as a reinforcing consequence for regular attendance at school, even though such might have a relatively high probability at a given moment, is not considered advisable. Dismissal from school may be reinforcing (not punishing) for truants; detention is also likely to be ineffective since it is a direct pairing of a noxious event with presence at school. Counter-conditioning is probably the procedure of choice.²

² An unpublished study by Schwitzgebel, Faltico, and Feeney (1967) selected 10 junior high school students from an underprivileged urban area who had never attended all 5 days of school in a given week for a period of 2 years. Home visits and small gifts

Arranging Contingencies

In our demonstration project, the basic wage of \$1.00 an hour was supplemented by positive consequences contingent on dependable and prompt attendance. The first problem, of course, is just to get the delinquent into the office under what he views as favorable conditions. The following conversation is a sample street-corner contact. In this particular instance, two Es saw five prospective Ss standing in a row in front of a dime store. It was mid-winter. The boys were huddled in their jackets; there was no conversation.

E: (approaching Ss) We're looking for some real people.

S1: That's us!

S2: See, look we move! (Waves arms; all laugh.)

E: Hey, you really are! (Exaggeratedly imitates arm waving, more laughing.)

We're from college and trying to find out what kids think about cops and teachers and school and stuff like that. Instead of reading a bunch of books, we decided just to go out and ask the kids themselves. We usually pay \$1 an hour or a little more for kids to talk into a tape recorder. . . . We'll pay you a dollar if you want to come with us to some restaurant around here; you can get some food too. You don't have to say anything you don't want to—no names or places or stuff like that. We're straight, and we're not cops.

S3: There's a pizza place down there (pointing).

S2: You mean we get a buck and mooch some free food?

E: Ya.

S2: (Offering E some peanuts from a small crumpled bag) Want some of these?

E: No thanks.

S4: I can tell you that the cops in (names a town) are a lot worse than those in (names another nearby town).

E: We want kids who know what they're talking about—not just a bunch of bull. You've got to have a record or know what you're talking about.

S2: Ha! I'm eligible all right! (Laughs) Here's my probation card. (Pulls out a partly torn probation card from his pocket and hands it to the E. E reads some of the rules aloud, such as "not keeping bad company"; group laughs.)

E: Anybody else got a record to qualify for the

contingent on gradual improvement of attendance showed significant results over a 6-week period. Due to the ending of the school year, no data were obtained regarding the maintenance of such behavior following termination of reinforcement. It might be noted, however, that in contrast to school attendance officers, who typically apply aversive sanctions following undesired behavior, the experimental procedure produced at least temporary gain. (One parent of Spanish-American descent commented to a graduate assistant that he was the first white person who had come to their house not to collect the rent or "tell us we done something wrong.")

restaurant? (Several boys point to a member of the group who reluctantly admits he does. Soon the *Es* and two boys go off to the pizza parlor. The *E* apologizes to the remaining three boys without records for not being able to take them along. Once the group gets to the restaurant, the boys ask if it would be all right just to order a coke and receive cash in place of food. The *E* agrees. Approximately 20 minutes is spent in the restaurant talking in a general way about police. The boys are then taken to the project office, play briefly with the tape recorder, are given a small cash bonus, and an appointment for another meeting is scheduled for the same time the following day.)

Obviously the delinquents are suspicious. They typically speculate that *E* is a gangster, policeman, homosexual, escaped mental patient, or—as one of them put it later—“just some new kind of nut.” The task of *E* is not to make guesses regarding the cognitions of the prospective patients, but to establish environmental conditions in which desired behavior will be prompted and reinforced. This is not excessively difficult. Delinquents often stand in collective boredom waiting for something to happen. The offer of a little free food, adventure, and money under a situation in which *Ss* outnumber *Es* in a safe public place has resulted in acceptance by an estimated 70% of eligible adolescents. It might be noted that food is used in this situation not only as a generalized reinforcer contingent on cooperative behavior but also as an unconditioned stimulus incompatible with anxiety (cf. desensitization procedures of Wolpe & Lazarus, 1966).

One group of nine *Ss* was then selected to receive reinforcers contingent on attendance, while a matched group of *Ss* served as controls by receiving bonuses contingent on an unrelated behavior. The *Ss* did not know when they would become eligible for a reinforcement or the nature of the reinforcement (i.e., a variable interval-variable ratio schedule).

In difficult cases when the patient does not have the necessary repertoire of social skills, it may be necessary to use a shaping procedure. In another study (Schwitzgebel, 1964), two boys were met at successively closer geographical locations to the office. The *E* arranged to meet the boys on the first two occasions at a downtown amusement center, the next meeting outside a subway station near the amusement center, then inside the subway station after the toll gate, eventually outside the subway station near the office, and finally at the office. Arrival at the designated location was casually rewarded with a cigarette or candy bar or perhaps money (in the form of a token put into the toll gate for *S*). The shaping required approximately six to eight meetings over a period of 2 weeks.

Recording Dependent Variable(s)

Both demonstration project groups improved in promptness and dependability of attendance over the first 20 sessions; however, the contingent-reinforcement group had a discrepancy of 11.2 minutes be-

tween appointment time and arrival time, while the noncontingent-reinforcement group showed a 15.6 minute discrepancy (Fisher test, $p < .025$). Since the noncontingent-reinforcement group received bonuses for other socially desirable behavior (e.g., putting coke bottles away at the end of the interview), generalization effects possibly obscured differences which might have occurred if more orthodox procedures had been used. Most *Ss* came to look upon the interviews as an interesting and challenging game; anticipatory approach behaviors became stabilized at a fairly high rate. Skinner (1953) and others have pointed out that gambling establishments use variable schedules of reinforcement. On one occasion, we noticed an *S* standing outside the office door, waiting for the chimes of a nearby church to ring, so that he could enter just at the appointed moment. Using the VI-VR schedule, nonverbal consequences could be faded out between the fifteenth and twenty-fifth session without noticeable decrement in promptness of attendance.

DISCUSSION

The strategy outlined here for introducing delinquents to psychotherapy is by no means entirely novel. Indeed, the implicit use of reinforcement techniques by therapists of widely divergent orientations has important implications since it suggests that an empirically based technology may begin to be established. In this way, the potential replicability of clinical procedures should be improved, and, at the same time, alternative methods should become more obvious. For example, Redl and Wineman (1952), in a chapter entitled “Techniques for the antiseptic manipulation of surface behavior,” note the ineffectiveness of making promises to behavior-disordered children and recommend that a group leader give “gratification grants without strings attached [p. 229].” They assert further that to tie these gratification grants to a condition would be “like promising a child cough medicine provided he goes without coughing for a day.” That is, they recommend the use of noncontingent consequences since they believe, apparently, that it is impossible to set appropriate criteria for future performance. What is not explicitly considered by Redl and Wineman, however, is making such grants contingent on modest *past* improvements of behavior (i.e., shaping).

Occasionally, clinicians will argue that “paying delinquents” is a form of bribery or a dishonest maneuver. Again, some ambiguity appears to be involved regarding very basic de-

tails of the reinforcement procedure. "Bribery" is usually defined as a promise of payment for a future performance. In instrumental conditioning, consequences follow desired behavior. On the variable schedule used in the study reported here, Ss were explicitly told that while they could depend on the base wage, bonuses were entirely the decision of E. Furthermore, the amount and frequency of reinforcers decreased over time in contrast to bribery situations where the victim usually demands increased payoffs as he becomes aware of the exploitation. Technically speaking, a promise to deliver a specified payoff following a specified performance is considered a "prompt." These may be useful in initiating behavior which is then subsequently reinforced by an operant procedure. Kubie (1958), apparently referring to the technique of prompting, has argued that a therapist may be required occasionally to use a "symbolic bribe."

We might also consider the frequently voiced objection to reinforcement procedures on the grounds that patients will become excessively dependent. This possibility exists, but its actual occurrence would likely be a result of poor therapeutic management. As previously mentioned, in our study, nonverbal consequences for attendance behavior were faded out between the fifteenth and twenty-fifth session. Extinction can be minimized, of course, by using variable schedules and gradually increasing the interval and ratio components.

In addition, other aspects of the therapy situation may take on secondary reinforcing properties, including the therapist himself. Therefore extrinsic consequences are not required. This situation may be particularly common in the case of therapists whom adolescents find "exciting" or "stimulating." Physiologically, one might attempt to measure such attributes by the occurrence in the patient of the "orienting reflex" (peripheral vasoconstriction, cephalic dilation, galvanic skin response, alpha desynchronization; see Sokolov, 1963) which seems to be both a prerequisite for learning and an intrinsically reinforcing event. The charismatic and colorful personalities of well-known therapists who work with behavior-disordered patients (cf.

Federn, 1962; Josselyn, 1957; Levenson, 1961) might be analyzed in terms of their ability to elicit orienting responses from patients who typically have a high threshold for arousal. One might even speculate further that the percentage of orienting and defensive reflexes during an initial interview with a given therapist could serve as a prognostic index.

This writer believes that a minimal technology now exists for introducing adolescent delinquents to psychotherapy or other "work" situations. Except for youths who persistently use narcotics, shaping dependable and prompt attendance of prospective patients has become for our experienced staff almost a routine matter. But just as any continuing social relationship tends to move from initial stereotyped patterns to those which are more subtle and unique, the preliminary socialization technique suggested here neither encompasses nor excludes the complexities and excitement of therapeutic interaction. It, at best, increases the probability that a continuing dialogue will occur.

REFERENCES

- ADDISON, R. M., & HOMME, L. E. The reinforcing event (RE) menu. *National Society for Programmed Instruction Journal*, 1966, 5, 8-9.
- AICHORN, A. *Wayward youth*. New York: Viking Press, 1935. (Reprinted New York: Meridian Books, 1955.)
- ESSLER, K. R. *Searchlights on delinquency*. New York: International Universities Press, 1949.
- FEDERN, E. The therapeutic personality as illustrated by Paul Federn and August Aichhorn. *The Psychiatric Quarterly* (New York State Department of Mental Hygiene, Utica), 1962, 1-15.
- FREUD, A. Adolescence. *The Psychoanalytic Study of the Child*, 1958, 13, 255-277.
- GORDON, J. E. Project Cause, the federal anti-poverty program, and some implications of subprofessional training. *American Psychologist*, 1965, 20, 334-342.
- HAMMAR, S. L., & HOLTERMAN, V. L. Referring adolescents for psychotherapy. *Clinical Pediatrics*, 1965, 4, 462-467.
- HOLLINGSHEAD, A. B., & REDLICH, F. C. *Social class and mental illness*. New York: Wiley, 1958.
- JOSSLYN, I. M. Psychotherapy of adolescents at the level of private practice. In B. B. Balser (Ed.), *Psychotherapy of the adolescent*. New York: International Universities Press, 1957.
- KUBIE, L. S. *Neurotic distortion of the creative process*. Lawrence, Kansas: University of Kansas Press, 1958.
- LEVENSON, E. A. Jam tomorrow—jam yesterday. *Etc.*, 1961, 18, 167-178.

- NEILL, A. S. *Summerhill: A radical approach to education*. London: Victor Gollancz, 1962.
- PELPZ, W. L. Psychotherapy of adolescents in private practice. In B. H. Balser (Ed.), *Psychotherapy of the adolescent*. New York: International Universities Press, 1957.
- PREMACK, D. Rate differential reinforcement in monkey manipulation. *Journal of the Experimental Analysis of Behavior*, 1963, 6, 81-89.
- PREMACK, D. Reinforcement theory. *Nebraska Symposium on Motivation*, 1965, 13, 123-180.
- REDL, S., & WINEMAN, D. *Controls from within*. Glencoe, Ill.: The Free Press, 1952.
- RIESSMAN, F., COHEN, J., & PEARL, A. (Eds.) *Mental health of the poor*. New York: The Free Press of Glencoe, 1964.
- ROSEN, B., BAHN, A. K., SHELLLOW, R., & BOWER, E. M. Adolescent patients served in outpatient psychiatric clinics. *American Journal of Public Health*, 1965, 55, 1563-1577.
- SCHMIDBERG, M. The analytic treatment of major criminals: Therapeutic results and technical problems. In K. Eisler (Ed.), *Searchlights on delinquency*. New York: International Universities Press, 1959.
- SCHOFFIELD, W. *Psychotherapy: The purchase of friendship*. Englewood Cliffs, New Jersey: Prentice-Hall, 1964.
- SCHWITZGEBEL, R. K. *Street corner research: An experimental approach to the juvenile delinquent*. Cambridge, Mass.: Harvard University Press, 1964.
- SCHWITZGEBEL, R. L. Short-term operant conditioning of adolescent offenders on socially relevant variables. *Journal of Abnormal Psychology*, 1967, 72, 134-142.
- SCHWITZGEBEL, R. L., FALTICO, G., & FEENEY, S. Instrumental conditioning of school attendance. Unpublished manuscript, Department of Psychology, University of California, Los Angeles, 1967.
- SKINNER, B. F. *Science and human behavior*. New York: Macmillan, 1953.
- SLAVSON, S. R. *The fields of group psychotherapy*. New York: International Universities Press, 1956.
- SOKOLOV, E. N. *Perception and the conditioned reflex*. New York: Macmillan, 1963.
- STIEPER, D. R., & WIENER, D. N. *Dimensions of psychotherapy: An experimental and clinical approach*. Chicago: Adline, 1965.
- WOLFE, J., & LAZARUS, A. A. *Behavior therapy techniques*. New York: Pergamon Press, 1966.

(Received March 5, 1968)

(Continued from page 70)

- ative Attitudes: Stephen M. Pittel* and Gerald A. Mendelsohn: Department of Psychiatry, Mount Zion Hospital and Medical Center, 1600 Divisadero Street, San Francisco, California 94115.
- Extreme Response Tendency as a Function of Emotional Adjustment and Stimulus Ambiguity: Russell P. Norman*: Department of Psychology, College of William and Mary, Williamsburg, Virginia 23185.
- Personality Factors Related to Response to Social Isolation and Confinement: Delmas A. Taylor*, Irwin Altman, Ladd Wheeler, and Estelle N. Kushner: NMRI-032-1h, National Naval Medical Center, Bethesda, Maryland 20014.
- An Attempt to Test Assumptions about Some Indications of Negativism on Psychological Tests: Elizabeth Fox and Sidney J. Blatt*: Department of Psychiatry, Yale University, 333 Cedar Street, New Haven, Connecticut.
- * Asterisk indicates author for whom address is supplied.
- Effects of Presession Information on Perception of the Therapist and Receptivity to Influence in a Psychotherapy Analogue: Roger P. Greenberg*: Department of Psychiatry, SUNY at Syracuse, 750 East Adams Street, Syracuse New York 13210.
- Digit Span: An Indicant of Trait or State Anxiety? William F. Hodges* and Charles D. Spielberger: Department of Psychology, University of Colorado, Boulder, Colorado 80302.
- The Client-Therapist Relationship Factor in Systematic Desensitization: Jon E. Krapfl and M. Mike Nawas*: Department of Psychology, University of Missouri, Columbia, Missouri.
- Effects of Vicarious Therapy Pretraining and Alternate Sessions on Outcome in Group Psychotherapy with Outpatients: Charles B. Truax and Donald G. Wargo*: Arkansas Rehabilitation Research and Training Center, University of Arkansas, Fayetteville, Arkansas 72703.

(Continued on page 95)

MEASUREMENT OF SOCIAL CAUTION: SELF-APPRAISAL, ROLE PLAYING, AND DISCUSSION BEHAVIOR

JAY S. EFRAN AND PAUL R. KORN¹

University of Rochester

3 experiments indicate some characteristics of socially cautious individuals, selected on the basis of a modified leaderless group discussion. As predicted from social learning theory, the cautious participants reported lower expectations for social success, were more reluctant to engage in other social-verbal tasks, were uncomfortable and inhibited at role playing, and indicated lack of confidence on an inventory of social behavior. Their expectations for other kinds of tasks were not different and they did not perform differently on a concept attainment task or in a conformity situation. Their responses suggested that the "caution" measured in the discussion procedure relates to peer rather than authority approval.

Some measure of the number of individuals who experience "shyness" as a major personal problem can be found in the sales records of books like *How to Win Friends and Influence People*, which promise relief from social discomfort. The distress such individuals experience can be acute; many learn to avoid social interaction whenever possible and find themselves trembling or "tongue-tied" when forced to participate. Because of the surplus meanings of terms that have been used to describe this pattern of behavior (e.g., introversion), the more neutral term "social caution" will be used in this analysis.

Rotter's (1954) social learning theory suggests that a person's reluctance to participate in social situations can be understood in terms of (a) the importance he attaches to approval, and (b) his expectations that given behaviors will lead to it. An individual who needs approval but lacks confidence that his behavior will elicit it may adopt an "avoidant" strategy; that is, faced with uncertainty about outcome, with stakes too high to risk experimentation, he hesitates, fails to respond, or withdraws. Since avoidant behavior is more probable when a situation involves negative reinforcement (e.g., disap-

proval), it may be more precise to describe the "shy" individual as having strong approval needs and reciprocally strong fears of disapproval. (His expectations for the former would be relatively low, and his expectations for the latter, relatively high). The distinction between positive and negative reinforcement is of importance because an individual who desires approval but who only minimally fears disapproval may display alternatives to avoidant behavior, such as vociferous attention-getting behavior. While this seems like an infrequent pattern, it does occur and one can envision a background of learning experiences to account for its development.

When the socially cautious person does participate, the content of his contributions is generally "safe." He waits until he has learned what kind of comment his "audience" will appreciate, or restricts his remarks to the patently nonoffensive. Ironically, either of these styles may, in itself, irritate listeners and bring about the disapproval the person was trying so hard to avoid.

This study investigates the generality of socially cautious behavior as measured by performance in a group discussion situation. Evidence is provided about the expectations and approval needs of cautious individuals.

The behavior of discussion participants was observed during a role-playing task, a concept attainment task, a conformity task, and a party-stunt choice task. In addition, Ss' self-evaluations were obtained on an "expectancy" questionnaire and the Minnesota In-

¹Now at the University of Connecticut. The authors wish to thank Edwin H. Lieberman for aid in the analysis of data and preparation of the manuscript. In addition to the authors, a number of individuals served as Es: Paul Stein and Robert Lewy in Experiment I, Harriet M. Braunstein in Experiment II, and R. Paul McNeal in Experiment III. Carl Spoto helped schedule S appointments.

ventory of Social Behavior (Williamson & Darley, 1937). It was hypothesized that cautious discussion participants would be less willing to engage in other kinds of social-evaluative performance, and would report lower expectations for success in social situations. In addition, they were expected to show signs of discomfort during role playing, to perform more poorly on the concept attainment task (especially with an "audience" present), and to agree more often with a confederate in the conformity situation.

EXPERIMENT I²

Subjects

Ninety-nine male college students, enrolled in an introductory psychology course, served as Ss. They participated in discussion groups of four to seven members. Ninety-three Ss returned individually for a second session.

Procedure

The Ss were seated around a table and given copies of a list of 16 tasks. They were told that they would each have to do *one* task in front of the group, and were led to believe that, within the limits set by the experimental design, task assignments would be based on their personal preferences. The five statements on the preference scale ranged from "1—These tasks sound like they might be fun or interesting" to "5—These tasks are objectionable enough so that I would *decline* doing them." The tasks were adaptations of party stunts, tests, and other common experiences involving various degrees of physical effort, social-evaluative performance, and unpleasant sensations. Sample items are: "Swallow one teaspoon of soapy water." "Talk to the group for approximately five minutes about your goals in life." "Blow up a balloon until it bursts." "Spend about five minutes telling the group a dramatic story about a college student."³

After Ss indicated their task preferences, one of the two Es collected the lists and left the room to "assign tasks." The Ss were then told to spend 25 minutes "trying to get to know each other." This precipitated

a leaderless group discussion, during which the second E served as an observer. When the time was up, group members evaluated each other's discussion performance on a sociometric nomination form. They had an opportunity to nominate individuals as having "participated actively in the discussion," having "participated comparatively little," and having "seemed rather shy." They could nominate as few or as many individuals in each category as they desired. Individuals identified each other by number cards fastened to the table in front of each place. The E, using a similar form, also evaluated the group members.

When the sociometric forms were collected, the first E returned and assigned tasks which Ss then performed. Task performance was not evaluated; the purpose of this part of the procedure was to insure that later Ss did not learn, via the grapevine, that task preferences were inconsequential. The Ss were always assigned preferred tasks. Before leaving, they were cautioned to maintain security about the procedures, and were invited to volunteer, on an individual basis, for "another experiment" on "concept formation." Almost all Ss did volunteer and individual appointments were scheduled.

When each S returned for his individual appointment, a concept attainment task (CAT) developed by Bruner, Goodnow, and Austin (1956) and modified by Weick (1964) was administered. A description of this task by Marcia (1966, p. 554), states that it

requires the subject to arrive at a certain combination of attributes of cards. The subject may eliminate certain attributes by asking whether a card is positive or negative for the concept and he may guess the concept at any time. He is penalized 5 points for every request, 10 points for every guess, and 5 points for every 30 seconds that passes before he attains the concept.

Half the Ss worked on this task in the presence of a confederate⁴ in addition to E. The purpose of this condition was to arouse additional evaluative stress. The confederate played the part of an S who was "running overtime" and had some forms to complete. To "keep things on schedule" E suggested that S get started while the confederate finished his forms. This provided an excuse for the confederate, who presumably had just completed the CAT, to stay in the room and watch S attempt the task. Shortly after S began, the confederate pretended to finish his forms and he then watched S's performance with considerable interest. After the administration of the CAT the confederate was dismissed. The Ss under both conditions next completed an expectancy question-

² Some of the data of this experiment were reported in a paper read at the Eastern Psychological Association Convention, New York, April 1967.

³ Copies of the instructions and forms used for these experiments can be obtained without charge from Jay S. Efran, University of Rochester, Rochester, New York, 14627, or for a fee from the National Auxiliary Publications Service. Order NAPS Document 00122 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001; remitting \$1.00 for microfiche or \$3.00 for photocopies.

⁴ Bruce G. Birnbaum, Kenneth H. Cowan, Bernard Demka, Eric L. Glazer, Robert Hegan, David C. Hewitt, William A. Hirsh, George McNaughton, Stephen F. Mendel, William R. Peters, Richard L. Reiter, David Ross, Joseph P. Sherer, Lee L. Suskin, and Daniel F. White served as confederates for Experiment I or Experiment III.

naire (EQ) listing 44 skills and abilities and requiring Ss to indicate how they compared in these areas with college students their own age. For each item they circled the percentage (0, 1, 4, 11, 23, 40, 60, 77, 89, 96, or 99) of students toward whom they felt superior. The items dealt with activities such as "singing," "getting along with teachers," "wood-working," "mathematics," etc.

For most Ss, scores on the Minnesota Inventory of Social Behavior (MISB)—a self-report measure of social anxiety—were available from an earlier (unrelated) group testing. The Ss who had been absent for that testing were asked to complete this form before leaving.

Results

Twenty-eight Ss were designated as active participants (AP) because they were nominated as "active" by all group members and the observer, and because no one considered them "shy" or having "participated little." Twenty Ss were classified as "socially cautious" (SC) because they received 20% or fewer nominations as "active" and 60% or more nominations in the "participated little" category. Twelve of these 19 Ss were viewed as minimal participants by *all* the raters. Most were considered "shy" by a large proportion of the group, and none received fewer than 40% nominations in this category. A conservative estimate of the reliability of the nomination form was obtained by comparing the nominations of the group members with those of the observer. The biserial correlations ($N = 93, r = .71, .80, .56$) for the three questions on which the classification of Ss was based indicate satisfactory reliability.

There was a significant difference between AP and SC Ss' preferences for the two tasks (combined) which directly involved verbal, social performance: telling the group a story and talking to them about goals in life ($t = 2.69, df = 46, p < .01$). The groups did not differ in their preferences for tasks which mainly required physical or intellectual effort or involved general aversiveness.

On an intuitive basis, six EQ items were expected to differentiate between AP and SC Ss. Differences significant at the .05 level or better were found for four of these items: making conversation ($t = 3.58$), getting along easily with most people ($t = 2.60$), leading groups ($t = 2.57$), social skills, generally ($t = 3.08$). Near significant differences were

found for the other two: making male friends ($t = 2.01$), and making a good impression at social functions ($t = 1.81$). Two items involving social and/or verbal skills, but not as closely related to the discussion task as the above items, produced significant differences: making female friends ($t = 2.47$) and directing plays ($t = 2.12$). In each case, SC Ss reported lower expectations for success than AP Ss. While the groups clearly differed in expectations for social success with peers, items concerning parents and other authority adults did not produce differences, nor did items concerning a wide variety of skills and abilities not directly related to social and verbal competence.

The scores of AP and SC Ss on the MISB, which has face validity for the measurement of lack of confidence in social situations, yielded a significant difference in the expected direction ($t = 2.50, df = 41, p < .05$).

CAT performance was unrelated to any of the other variables of the study and the presence or absence of the confederate had no discernible effect on adequacy of performance.

EXPERIMENT II

Subjects

The AP and SC Ss of Experiment I were telephoned a semester later and invited to participate individually in this experiment. Because of scheduling difficulties, only 22 AP and 18 SC Ss participated.

Procedure

A female *E* read the following instructions:

I am going to ask you to carry on several imaginary phone conversations using this prop telephone. As you can see, the phone isn't connected. As you pretend to be talking to someone, I'll just tape-record your responses. As a kind of warm-up exercise, I'd like you to pretend to call the information operator to get a phone number for this person. [A card with the name "Bill Dudley" was shown to the subject.] Any questions? Now remember, you are to engage in a conversation. Okay, you can begin now.

The following additional role-play situations were then given to S:

1. You have discovered a hole in a suit you just brought back from the cleaners—you are sure it wasn't there when you took the suit in. Call the cleaners about it.
2. Your friend's grandmother died. Call to offer your condolences.

3. Call to make a blind date with a girl whose phone number was given to you by a friend.

4. A friend of yours has divulged something to others which you told him in confidence. Call him about it.

5. Today is the birthday of a friend who lives in California. Call him as a birthday surprise.

6. You have made plans for a special date. Just before leaving to pick up your date, she calls and says that she can't go because she has to stay home with her sick brother. You know that she does not have a brother!

The *Ss* performed each situation once except for the conversation with the dry cleaner, which they were asked to do again as angrily as possible. After each role-play, *Ss* indicated how "comfortable" they had felt on a 7-point scale ranging from "completely at ease" to "very ill-at-ease." At the same time, *E* rated *Ss*' apparent anxiety level on a 5-point scale. Following the role-playing, *Ss* were asked to imitate the sounds of six animals: a cat, dog, cow, duck, horse, and lion. These imitations were also recorded.

Results

An independent rater, who did not know the identity of the *Ss*, rated the taped role playing and the imitations.⁵ A 4-point scale of "cautious" performance was used, ranging from "an inhibited, cautious, self-conscious attempt" to "an uninhibited, enthusiastic attempt." In making these judgments, the rater tried to avoid evaluating role-playing skill per se—he evaluated the caution or freedom with which *S* approached the task. Obviously, skill and caution are not orthogonal dimensions, but the focus was on the latter. The reliability of ratings, assessed by having a second rater make the same judgments, was satisfactory for role playing ($r = .70$), but low for the animal imitations ($r = .41$).⁶ The *Ss* who were rated less inhibited in role playing reported greater subjective comfort ($r = .38$, $p < .05$) and were rated as less anxious by *E* ($r = .53$, $p < .01$). The *Ss* who seemed less inhibited in the animal imitation task reported greater comfort ($r = .46$, $p < .01$), but were not seen as significantly less anxious by *E* ($r = .21$, *ns*). The animal imitation and role-playing ratings correlated significantly with each other ($r = .34$, $p < .05$), as did the comfort and anxiety ratings ($r = .47$, $p < .01$).

⁵ Andrew Hoffman served as rater.

⁶ One *S* was dropped from this analysis because of tape transcription difficulties.

As predicted, the role playing of AP and SC *Ss* differed. Relevant biserial correlations and tests are presented in Table 1. Animal imitation performance was also rated significantly different ($t = 2.48$, $p < .05$), but proved a less powerful discriminator.

The interpretation of relationships between role playing and the self-report measures of Experiment I must be tempered by the fact that the *S* sample was not random, but had been preselected on the basis of discussion performance. However, it is of interest that role playing relates to MISB scores ($r = .36$, $p < .05$), to preference for social-verbal evaluative tasks ($r = .49$, $p < .01$), and to the sum of social-verbal EQ items ($r = .53$, $p < .01$) as well as the individual items. As before, no significant relationships were found with task and EQ items less directly social or verbal in nature.

EXPERIMENT III

Subjects

A new group of 79 male *Ss* enrolled in an introductory psychology course participated in discussion groups of four to seven members. The procedure for the selection session was the same as that used in Experiment I, except that slightly modified task preference and sociometric forms were used. On the basis of this procedure, 24 *Ss* were designated as AP and 18 were classified as SC. These *Ss* were then telephoned by a different *E*, and 16 AP and 9 SC *Ss* agreed to participate in an experiment on

TABLE 1
RATINGS OF ROLE PLAYING FOR ACTIVE AND SOCIALLY CAUTIOUS DISCUSSION PARTICIPANTS

Rating	<i>M</i>	<i>SD</i>	<i>t</i>	Biserial r
Anxiety observed by <i>E</i>				
Active participants	29.00 ^a	4.38	4.60*	.60*
Cautious participants	21.22	6.02		
Comfort reported by <i>S</i>				
Active participants	44.14 ^b	5.14	3.85*	.53*
Cautious participants	37.83	4.90		
Inhibition evaluated from tape recordings				
Active participants	12.55 ^c	2.66	5.00*	.63*
Cautious participants	7.94	3.01		

Note.—For active participants, $N = 22$; for cautious participants, $N = 18$.

* $p < .001$.

^a A higher score indicates less anxiety.

^b A higher score indicates greater comfort.

^c A higher score indicates less inhibition.

"taste perception." They were unaware of any connection between this request and their previous participation.

Procedure

The "taste perception" session used a conformity procedure developed by Rettig and Sinha (1966): A few drops from each of 16 bottles were placed on the tongues of an *S* and a confederate to test "taste thresholds." *Ss* were told that each sample would be sweet, sour, salty, or bitter, but actually only four bottles contained flavoring (sugar, salt, lemon juice, and quinine sulphate). The rest contained water and food coloring. On some trials *S* responded first, and on others the confederate responded first. The confederate's responses were, of course, predetermined. After this taste test, the *S* and the confederate completed forms which asked about their confidence in making judgments, their feelings about the independence of their judgments, and their perception of the purpose of the experiment.

Results

As in Experiment I, the 24 AP and 18 SC *Ss* differed in their responses to the combined social-verbal evaluative items on the task preference form ($t = 3.01, p < .01$).⁷ Again, no other items were significantly related to the sociometric classification.

It was predicted that SC *Ss*, compared with AP *Ss*, would be more influenced by the confederate's responses. Neither the scores indicating conformity to the confederate's responses nor those indicating taste accuracy reveal differences between the groups. On the posttest questionnaire there was a tendency for the SC *Ss* to report that the "other *Ss*'s" responses had a greater effect on them, but this finding did not reach significance with the small sample used ($t = 1.94, df = 23, p < .10$).

DISCUSSION

Some characteristics of active and cautious discussion participants can now be described. In accord with social learning theory predictions, cautious participants acknowledge lower expectations for success in a variety of social and verbal skills. This does not reflect a general questionnaire response bias

since their stated expectancies for other skills and abilities are the same as those of active participants. Their lack of social confidence is also seen in their response to the MISB. Moreover, the significant difference between the scores of active and cautious participants on this instrument serves as an indication of its validity.

The generality of the behavior tapped by the sociometric classification used in these studies is evidenced by the replicated difference in the willingness of active and cautious individuals to perform other social-verbal tasks and by the differences in their role playing. These latter differences are particularly striking: Cautious discussants apparently look, feel, and sound more inhibited and anxious when role playing. Further, this relationship between discussion behavior and role playing was found despite the fact that the tasks involved somewhat different skills, and were administered a semester apart in different settings (one individually and one in a group) by *Es* of different sexes.

The results just discussed indicate that the AP-SC classification measures a characteristic more general than discussion behavior in a male group. However, the experiments also provide data indicating some of the limits of this generality. The hypotheses that SC *Ss* would perform more poorly on the CAT, would be more sensitive to "audience" pressure on this task, and would go along with the confederate's choices more often on the "taste test," were all based on the assumption that a cautious style in the discussion situation would have very general implications for performance on other kinds of tasks. None of these hypotheses was supported. The most parsimonious interpretation of this outcome is that the *Ss* we identified as "cautious" had in common only a lack of confidence concerning tasks construed as primarily social and/or verbal. This is, of course, what the *Ss* told us on the EQ: they did not report lower expectations for success in intellectual, athletic, mechanical, or artistic endeavors. Therefore, the CAT task, which was presented as intellectual, and the "taste test," which was presented as a psychophysical investigation, did not elicit "cautious" responses. Perhaps the difficulties investigators have had in finding

⁷ The revised task preference form used in this study contained three social-verbal evaluative items: telling the group a story, talking to them about goals in life, and portraying a master of ceremonies introducing a juggling act.

personality correlates of "conformity" behavior relates to the common practice of defining conformity tasks as some kind of "perceptual judgment" situation.

An examination of the EQ items which do and do not differentiate between AP and SC Ss indicates that while the groups generally respond differently to items which describe social skills, they do not respond differently to items which focus on relationships with parents, teachers, employers, or other adults of authority. Thus it would seem that the kind of social caution measured in the discussion situation relates primarily to the attainment of peer reinforcement rather than reinforcement from authority figures. These findings concerning generality suggest that the class of situation and the reinforcing agent must be specified to obtain reasonable prediction. Basing measurement techniques on broad concepts such as "self-esteem" or "self-concept," where these factors are not specified, should continue to lead to low-level prediction.

Theoretically, Ss classified as "cautious" in this series of experiments might have had low expectations for social success but attached little importance to obtaining social reinforcements. If this were so, the Ss selected would not be the type proposed in the introduction, where the "shy" person was described as attaching importance to the receipt of approval. However, Experiment II provides evidence on this point. When forced to role play (by the demands of the experimental situation), SC Ss stated that they experienced discomfort and *E* reported congruent symptoms of the Ss' distress (e.g., sweating, stammering, tremor, etc.). Together

with our other measurements, these manifestations of evaluative concern clearly indicate that SC Ss were not indifferent to the social reinforcements of the situation, and that they are more accurately described as "fearful of appearing foolish" than "disinterested in social approbation."

The leaderless group-discussion technique differed from those often used (e.g., Bass, 1954) in that no specific topic was assigned. The Ss were merely instructed to "get to know one another." It was felt that this definition of the task would facilitate its perception as "social" rather than "intellectual," and would help prevent Ss from beginning at a disadvantage because they had not read recent newspapers or were not involved with a particular campus issue.

REFERENCES

- BASS, B. M. The leaderless group discussion. *Psychological Bulletin*, 1954, 51, 465-492.
- BRUNER, J. S., GOODNOW, I. J., & AUSTIN, G. A. *A study of thinking*. New York: Wiley, 1956.
- MARCIA, J. E. Development and validation of ego identity status. *Journal of Personality and Social Psychology*, 1966, 3, 551-558.
- RETTIG, S., & SINHA, J. B. P. Bad faith and ethical risk sensitivity. *Journal of Personality*, 1966, 34, 275-286.
- ROTTER, J. B. *Social learning and clinical psychology*. Englewood Cliffs, N. J.: Prentice-Hall, 1954.
- WEICK, K. E. Reduction of cognitive dissonance through task enhancement and effort expenditure. *Journal of Abnormal and Social Psychology*, 1964, 68, 533-539.
- WILLIAMSON, E. G., & DARLEY, J. G. *The Minnesota Inventory of Social Behavior*. New York: Psychological Corporation, 1937.

(Received March 8, 1968)

SELF-ESTEEM:

A SELF-SOCIAL CONSTRUCT¹

ROBERT C. ZILLER, JOAN HAGEY, MARY DELL C. SMITH

University of Oregon

AND BARBARA H. LONG

Goucher College

Self-esteem is defined within a context of self-other orientation, and an instrument is described which is designed to measure the evaluation of the self in relation to significant others using topological representations of self and others and involving limited verbal demands. The results of a program of research are described which emanate from an evolving theory of social self-esteem. The results suggest that self-acceptance and social acceptance are inextricably combined and raise serious doubts about the meaning of earlier results concerning self-esteem which were based upon verbal self-reports.

Self-esteem is usually defined as the individual's perception of his worth. In evaluating the self, however, few physical cues are available which provide a reliable basis for an estimate. Still, the individual has recourse to paired comparisons of the self and significant others; that is, self-evaluation evolves in terms of social reality (Festinger, 1954). Self-evaluation, then, emerges largely within a social frame of reference.

According to this definition of self-esteem, if the social environment changes, a corresponding change in self-esteem may be anticipated. It is now proposed that the person's response to the social environment is a function of self-esteem. Self-esteem mediates social stimuli and response (Social stimuli \rightarrow Self-esteem \rightarrow Response). It is proposed here that self-esteem is a component of the self system which regulates the extent to which the self system is maintained under conditions of strain, such as during the processing of new information concerning the self. Thus, for example, evaluations of either a positive or negative nature do not evoke immediate, cor-

responding action by the individual with high self-esteem. New information is examined on the basis of its relevance and meaning for the self system and is disregarded if its meaning tends to be tangential. In this way the organism is somewhat insulated from the environment or is not completely subject to momentary environmental contingencies.

Persons with low self-esteem, on the other hand, do not possess a well-developed conceptual buffer for evaluative stimuli. In Witkin's terms (Witkin, Dyk, Fotherson, Good-enough, & Karp, 1962), the person with low self-esteem is field dependent; that is, he tends to passively conform to the influence of the prevailing field or context. Since the individual's behavior is directly linked to immediate environmental circumstances and is not mediated or differentiated and integrated by the self-concept, he is thereby inclined toward inconsistency.

Thus, the concept of self-esteem as described here is linked to the concept of personality integration as used by Lewin (1935). Development was described by Lewin as including an increase in the number of the relatively independent subparts of the person (differentiation) and increasing the unity of the person (integration or organization). Similarly, Piaget (1947) presents the concepts of assimilation and accommodation. It is anticipated that the behavior of persons with high self-esteem is more integrated and that their cognitive processes are characterized by a

¹The research program from which this report emanated was supported in part by a grant to the senior author by the National Science Foundation and in part by the United States Office of Education through a contract to the Center for the Advanced Study of Educational Administration, University of Oregon.

²Requests for reprints should be sent to Robert C. Ziller, Department of Psychology, University of Oregon, Eugene, Oregon 97403.

selective consideration of relevant social elements (stimulus control).

COMMUNICATION AND THE MEASUREMENT OF SELF-ESTEEM

Previous research concerning self-esteem has not emphasized sufficiently the social nature of the self system. The failure to incorporate and weight social factors within the self-evaluation framework may have contributed, in part at least, to the disappointing state of the investigation of self-esteem.

A second shortcoming of earlier studies in this area is their descriptive nature, which, coupled with the serious shortcomings of the measurement techniques, have left the area at a low level of theoretical development.

Finally, and most seriously, previous research has largely involved a verbal self-report measure of self-esteem. Kelly (1955) qualifies his assertions continuously by pointing to the most tentative of his assumptions, that S's word labels for his constructs mean what the examiner thinks they mean. Finally, Kelly suggests that if a test

can be arranged to produce a kind of protocol which can be subjected to a meaningful analysis, independent of words, we shall have made progress toward a better understanding of the client's personal constructs [p. 268].

The approach used here involves a method of communication with limited verbal demands and relies primarily upon rudimentary forms of abstraction which are assumed to predate verbal communication systems.

The approach is a confluence of the approaches of DeSoto and Kueth (1959), Kelly (1955), and an evolving theory of self-other orientation. The measurement approach assumes that the human organism finds it expedient to order and categorize or to structure generally the multitude of self-surrounding stimuli. The processes used by the respondents are expected to be somewhat idiosyncratic, but owing to commonality among human experience, sensory processes, and classification systems, the evolving abstraction systems possess sufficient commonality that the basis of a communication system exists. Some of these processes include extent of separation between objects (Kueth, 1962), number of objects in a category, and

ordering of objects (DeSoto, London, & Handel, 1965).

The most relevant ordering process with regard to self-esteem is what DeSoto, London, and Handel refer to as "spatial paralogic" and "linear ordering." It is observed that people are prone to place elements in a linear ordering to the exclusion of other structures, and that they handle linear ordering more easily than most other structures (Coombs, Raiffa, & Thrall, 1954; DeSoto, London, & Handel, 1965). Indeed, DeSoto, London, and Handel note that serial ordering proceeds more readily in a rightward direction than in a leftward direction. The tendency to attribute greater importance to the object placed at the extreme left position in a horizontal display has been noted by Morgan (1944).

The measure of self-esteem developed here utilized the serial ordering predilection of Ss within a social context. (See Figure 1.)

The measure involves presenting a horizontal array of circles and a list of significant others (including the self) such as those used by Kelly (1955). The task requires S to assign each person to a circle. The score is the weighted position of the self. In accordance with the cultural norm, positions to the left are assumed to be associated with higher self-esteem.

The item in Figure 1 is one of six items in the student form of the instrument. The other five self-esteem items involve the following sets of significant others: (a) doctor, father, friend, mother, yourself, teacher; (b) some-

DIRECTIONS. The circles below stand for people. Mark each circle with the letter standing for one of the people in the list. Do this in any way you like, but use each person only once and do not omit anyone.

- F—someone who is flunking
- H—the happiest person you know
- K—someone you know who is kind
- S—yourself
- Su—someone you know who is successful
- St—the strongest person you know



FIG. 1. Measure of self-esteem.

one you know who is a good athlete, someone you know who is a good dancer, someone you know who is funny, someone you know who gets good grades, yourself, someone you know who is unhappy; (c) an actor, your brother or someone who is most like a brother, your best friend, a dean of students, yourself, a salesman; (d) someone you know who is cruel, your grandmother, a housewife, a policeman, yourself, your sister or someone who is most like a sister; (e) doctor, father, friend, nurse, yourself, someone you know who is unsuccessful.

In a study involving 75 randomly selected students from Grades 7 through 12, the split-half reliability (odd-even) was .80 corrected for length (Long, Ziller, & Henderson, in press). Split-half reliability (odd-even) for the adult form^a was .85, uncorrected for length, in a study (Mossman & Ziller, 1968) involving 60 neuropsychiatric patients. Test-retest reliability for 86 sixth and seventh graders was .54 for the student form.

VALIDATION

The measure of self-esteem proposed here is assumed to involve social reasoning and a norm of hierarchical ordering of social objects in a horizontal line from left to right. This assumption was examined in a series of separate studies.

In the first of these (Ziller, Megas, & DeCencio, 1964), 45 patients in an acute neuropsychiatric treatment ward were presented with seven circular pieces of white felt cloth 2 inches in diameter. The circles

^a The six sets of social objects included in the adult form of the instrument are: (a) doctor, father, a friend, a nurse, yourself, someone you know who is unsuccessful; (b) doctor, father, friend, politician, yourself, an employer; (c) someone you know who is a good athlete, someone you know who is popular, someone you know who is funny, someone who knows a great deal, yourself, someone you know who is unhappy; (d) an actor, your brother or someone who is most like a brother, your best friend, yourself, a salesman, a politically active person; (e) someone you know who is cruel, a judge, a housewife, a policeman, yourself, your sister or someone who is most like a sister; (f) a defeated legislative candidate, the happiest person you know, someone you know who is kind, yourself, someone you know who is successful, the strongest person you know.

were marked by symbols indicating the person they represented. These persons included nurse (N), nurses' aide (NA), other patients in the ward (O), psychiatrist (Pi), psychologist (Po), social worker (SW), and yourself (Y). A list of the symbols and their referents were placed on a table in alphabetical order for the S's information. These were also read to the Ss. The Ss were instructed to arrange the circles on a black felt board, 2 x 2½ yards, in any way that they wished.

In support of the serial ordering tendencies proposed by DeSoto, London, and Handel, the majority of Ss arranged the symbolic circles in a straight line from left to right. By assigning weights to the left-right positions and calculating the mean weighting of the seven symbols, the resulting order of the symbolized positions was: psychiatrist, psychologist, social worker, nurse, nurses' aide, other patients, and yourself. It is apparent that a left to right status hierarchy of the social objects emerges.

A similar analysis was made of the left to right location of a low-status other person by college students using the student form (see Table 1). It was noted that the "unhappy" person was placed in the last position to the right 48% of the time; "someone you know who is unsuccessful," 56%; "someone you know who is cruel," 64%; and "someone you know who is flunking," 59%.

Further support for the assumption of a left to right paralogic was found in the association of the self-esteem score as previously derived and a second technique of scoring.

TABLE 1
DISTRIBUTION OF THE LOCATION OF THE "NEGATIVE SIGNIFICANT OTHERS" IN FOUR ITEMS OF THE SELF-ESTEEM MEASURES, HORIZONTAL ARRANGEMENT POSITIONS

Item	6	5	4	3	2	1
Unhappy	16	7	7	7	15	48
Unsuccessful	21	3	5	9	5	56
Cruel	27	2	4	1	3	64
Flunking	31	1	2	3	4	59

Note.—For Unhappy, N = 150; for Unsuccessful, N = 147; for Cruel, N = 154; for Flunking, N = 172. Percentage signs are omitted.

This involved the identification of the most negative, significant other for each set of significant others and calculating the distance, in number of circles, between the low-status other and the self. (Only four items which involved a clearly differentiated low-status other were included in this analysis. These items were *a*, *b*, *d*, and *e*.) This method of scoring was suggested when it was noted in Table 1 that the low-status other was sometimes located in the left position, indicating that arrangement of the self and others may be based on other than a left to right hierarchical ordering in some cases.

The correlation between the scores derived by these methods was .33, $p < .05$ ($N = 163$ male and female college students). Only the results with regard to males were statistically significant, however ($r = .46$, $N = 61$, $p < .001$; females, $r = .14$, $N = 102$, $p < .10$).

A similar analysis of the two scoring methods was made using the adult form with male neuropsychiatric patients (Mossman & Ziller, 1968). The results corroborate the findings with regard to the male college sample ($r = .56$, $N = 60$, $p < .001$).

In yet another study of the left-right serial ordering phenomenon, a children's form of the self-esteem measure was used (Henderson, Long, & Ziller, 1965). The analysis involved the responses of 48 boys and girls ranging in age from 7 to 14 who had applied for corrective training at a reading study center, plus 48 controls matched for age, sex, and general intelligence. The Ss were given a paper with a long horizontal line. They were next presented in random order six circles with pictures representing self, friend, and a "smart," "dumb," "funny," and "bad" classmate. The children were told to paste these symbolic circles in a row on the line. It was found that children placed the "smart" classmate to the left and a "bad" classmate to the right to a significant degree.

Evidence that the left-right serial ordering is not a phenomenon limited to persons within the United States is found in the analysis of the location of "someone you know who is unsuccessful" in the student form of Item *e*. With regard to 92 boys and girls from Form I of the M.V.D.M. High School in Vesakapatnam, Andhra, South India, the frequency with

which the "unsuccessful" person was located in the positions from *right to left* was 74, 9, 4, 0, 1, and 4. With regard to an American sample of 94 boys and girls, the corresponding frequencies were 53, 17, 15, 3, 3, and 3.

Another test of the left to right hierarchical ordering assumption was the association between the weighted position of "yourself" among five others including "someone you know who is flunking," "the happiest person you know," "someone you know who is kind," "someone you know who is successful," and "the strongest person you know" under conditions where the social objects were to be arranged horizontally as in Figure 1 and vertically. In the vertical display, the higher position of the self is assumed to represent higher self-esteem. The correlation between these two measures was .50 ($N = 82$, $p < .05$).

An analysis of the location of the lowest status other, "someone you know who is flunking," indicates (see Table 2) that the number of reversals in the placement of the low-status other is reduced in the vertical arrangement (9% vs. 31%). The vertical arrangement may introduce greater item visibility, however.

A third approach to the validation of the social self-esteem (SSE) measure was a correlational analysis of SSE with existing measures of the construct. The measures selected for comparison were those most frequently referenced in the literature (Wylie, 1961) and a more recent device developed for research purposes by Cutick (1962) and used by Diggory and her collaborators (Diggory-Farnham, 1964). With the exception of the

TABLE 2
DISTRIBUTION OF THE LOCATION OF "SOMEONE WHO IS FLUNKING" IN TWO IDENTICAL SELF-ESTEEM ITEMS (VERTICAL VS. HORIZONTAL ARRANGEMENTS)

Location arrangement	6	5	4	3	2	1
Vertical	9	.06	.06	3	3	84
Horizontal	31	1	2	3	4	59

Note.—In the vertical arrangement location "6" was the first or top position in the hierarchy. In the horizontal arrangement location "6" is the first position in the left-right hierarchy. Percentage signs are omitted.

SSE, all the measures were based on self-reports. Thus, the Bills-Vance-McLean (1951) Index of Adjustment and Values required *S* to rate himself with reference to each of 49 adjectives as to how often he was "this sort of person." Six-week test-retest reliability was .90 ($N = 100$).

Diggory's Self-Evaluation Questionnaire asks *S* the percentage of time that he expected to succeed in eight given situations. The reliability is not reported.

Coopersmith's (1959) Self-Esteem Inventory contains 54 items concerned with *S*'s perceptions in four areas: peers, parents, school, and self. The form was modified slightly to make it more appropriate for a college population. The Self-Esteem score is twice the sum of the high self-esteem items (as agreed upon by five psychologists) marked "like me" and low self-esteem items marked "unlike me." Reported test-retest reliability after 5 weeks was .88.

In an earlier test of the relationship between the SSE and a single item, overall self-evaluation (Ridgeway, 1965), a negative but not statistically significant relationship had been found ($r = -.15$, $N = 100$, $p < .5$). Given the different theoretical frameworks and method of communication upon which the measures are based, a significant relationship was not anticipated in the present study. The purpose of this study in the program was to establish the independence of the SSE more systematically.

The correlation matrix for these three scales and the SSE for each sex is shown in Table 3. None of the correlations with SSE

were statistically significant. Once again, sex differences are quite apparent in the inter-correlations among the measures. For male *Ss* ($N = 33$), the highest correlation, $r = .60$, was between Diggory's and Bills' measures. Significant correlations with regard to male *Ss* were also found between Diggory's and Coopersmith's measures ($r = .37$), and Bills' and Coopersmith's ($r = .46$). The only significant correlation found for female *Ss* ($N = 53$) was between Diggory's and Bills' measures ($r = .29$). These results are worrisome, even though they were anticipated. Yet, the results may be interpreted to indicate that the SSE and the other measures of self-esteem are in different psychological domains. The SSE in contrast to the other devices is a nonverbal, "low visibility" instrument, and also incorporates a social frame of reference.

One of the universal criticisms of the most frequently used measures of self-acceptance is that they are about equally correlated with socially desirable responses as they are with each other (Crowne, Stephens, & Kelly, 1961). For example, the greater the tendency to give socially desirable responses, the less the reported discrepancy between self and ideal self.

Using the Crowne-Marlowe (1964) measure of socially desirable response tendencies, and relating it to the SSE as well as Diggory's measure of self-evaluation, correlations of $-.36$ and $.65$ ($N = 24$, $p < .05$ for both) were found for sophomore female volunteers for an experiment. Higher self-esteem as measured by the Diggory device was associated with a tendency to give socially desirable re-

TABLE 3
INTERCORRELATION MATRIX OF FOUR MEASURES OF SELF-ESTEEM

Measure	(2)		(3)		(4)	
	Males	Females	Males	Females	Males	Females
Bills-Vance McLean Index (1)	.46**	.17	.60***	.29*	-.10	-.14
Coopersmith's Self-Esteem (2)			.37*	.23	.02	.04
Diggory's Self-Evaluation (3)					-.09	.21
Social Self-Esteem (4)						

Note.—For males, $N = 33$; for females, $N = 53$.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

sponses. The opposite relationship was found using the SSE.

SELF-ESTEEM AND SOCIAL ACCEPTANCE

Turning now to construct validation procedures, one of the earliest studies in the series examined the frequently hypothesized relationship between acceptance of self and acceptance by others (Mann, 1959; Wylie, 1961). In one of the reported studies, Cooper-smith (1959) found that fourth, fifth, and sixth graders showed a significant positive correlation (.34) between self-esteem and popularity. The rationale for the relationship is often tautological (see Rogers, 1951, p. 520) but the findings are nevertheless consistent. Within the present framework, self-acceptance and the acceptance by others are perceived as inextricable components of social self-esteem.

The Ss in this study (Ziller, Alexander, & Long, 1964) were 321 sixth-grade students in 11 classrooms from four elementary schools. The Ss were all white, and the composition of the classes remained unchanged throughout the school day. All Ss completed a sociometric item asking them to name the five children with whom they would most like to play. Twenty-five children (17 boys and 8 girls) who were unchosen and 25 children (17 boys and 8 girls) who were most highly chosen from the same classes as the unchosen were administered one item of the social self-esteem measure. The social set included "doctor," "father," "friend," "the person with whom you are most happy," "mother," "yourself," "the most successful person you know," and "the person with whom you are most comfortable." The directions were read to Ss. The mean position of the popular children (1 being the left position) was 3.8 and of the unpopular 5.7 ($t = 3.87$, $p < .005$).

SUCCESS AND FAILURE OF POLITICAL CANDIDACY AND SELF-ESTEEM

One of the difficulties of studying changes in the self-concept is that conditions associated with changes in the self-concept are not readily generated and are rarely encountered under circumstances amenable to statistical analysis. A political election, however, provides an exceptional opportunity to study

TABLE 4

CHANGES IN SELF-ESTEEM OF WINNING AND LOSING POLITICAL CANDIDATES

Candidates	Direction of change in self-esteem		
	Increased	Decreased	No change
Winners	15	4	4
Losers	8	11	2

$$\chi^2 = 6.01 (p < .05).$$

changes in the self-concept associated with winning as opposed to losing the election.

One month prior to the 1967 state election in Oregon, 44 candidates for the state legislature were administered the first four items of the adult Form II of the SSE. The same items were again administered to the same candidates approximately 1 month after the election. Each candidate was approached individually and completed the form in the presence of the data collector. The results are provided in Table 4. Fifteen of the 23 winning candidates increased in self-esteem as opposed to 8 of the 21 losing candidates. Moreover, 11 of the 21 losing candidates decreased in self-esteem as opposed to 4 of the 23 winners. The results are significant at the .05 level of confidence ($\chi^2 = 5.99$).

SELF-ESTEEM AND CONSISTENCY OF SOCIAL PARTICIPATION

In an experiment by Mossman and Ziller (1968), it was hypothesized that self-esteem is: (a) positively related to frequency of participation in group discussion, and (b) is associated with the organism's consistency of response. Relative to the first hypothesis, it has already been demonstrated that high self-esteem is associated with social acceptance. This finding suggests that in a group discussion the individual with high self-esteem will receive and expects to receive verbal and nonverbal cues from the other members which invite or support his bid for participation. In addition, the high self-esteem—high socially accepted member may be expected to receive more social reinforcement for his participation. Thus, the high self-esteem—high socially accepted member is assumed to receive more self-reinforcement and social rein-

forcement for participation in group discussion, which results in a higher level of participation.

Here, high self-esteem is assumed to be associated with a higher potential for self-reinforcement and a higher probability of social reinforcement. To some extent self-reinforcement and social reinforcement may be complementary. When social reinforcement is withheld, the individual with high self-esteem has recourse to self-reinforcement. The individual with low self-esteem is more dependent upon social reinforcement, leading, as we stated at the outset, to less stable participation.

Borgatta (1962, p. 256) demonstrated a significant correlation between Cattell's Guilt Proneness versus Confident Adequacy subtests and total activity in a group discussion ($r = -.34$, the more guilt, the less activity) and between Edwards' Abasement subtest and total activity ($r = -.30$, the less abasement, the more activity). The Ss in the study were 76 neuropsychiatric patients who were members of four "autonomous problem-solving groups" similar in purpose to those discussed by Fairweather (1964, p. 171). Each of the four groups was observed during one session a week over a 3-week period. The observer recorded the total amount of interaction units per individual. The adult forms of the SSE were administered at the end of the third session.

The self-esteem scores were ordered as anticipated with regard to levels of verbal participation: high interactors (those who contributed more than 5% of the interaction units across the three sessions), 21.6 (the sum of six ratings); low interactors (those who contributed between 2% and 5% of the interaction units, 19.8; and the noninteractors (those who contributed 1% or less of the interaction units across the three sessions), 17.4. Furthermore, the results were statistically significant ($F = 3.37$, $df = 1/57$, $p < .05$).

In order to test the second hypothesis, the variance of the relative frequency of interaction units across the three sessions was used as the measure of consistency of social behavior for each individual. A highly consistent interactor was defined as a group member

whose relative frequency of interaction variance was .03 or less. The .03 point of division created two equal-size categories of Ss with regard to consistency of verbal participation. However, only the consistency of high interactors was analyzed since consistency among low interactors and noninteractors would be a statistical artifact stemming from a ceiling effect. As hypothesized, there was a significant difference ($p < .05$) in self-esteem scores between the high interactors-low consistency (SSE = 15.90, $N = 10$) and high interactors-high consistency (SSE = 26.73, $N = 11$) categories of Ss.

SELF-ESTEEM AND SOCIOECONOMIC STATUS

Although Wylie's review of the literature contains no reference to an analysis of a relationship between socioeconomic status and self-esteem, more recently two studies (Coopersmith, 1967; Rosenberg, 1965) indicate a positive relationship, although only the results of the first are statistically significant. The rationale for the relationship is that social status is one of the most striking indexes of prestige and success. Persons higher in the social system have more prestigious occupations, have higher incomes, and tend to live in larger and more luxurious houses located in more desirable neighborhoods. These persons are perceived as more successful and tend to receive material and cultural benefits that might lead them to believe that they are generally more worthy than others.

Coopersmith (1967) points out, however, that children's social position emanates from experiences in school and the neighborhood rather than in an occupational context. These attenuating considerations notwithstanding, it is proposed children from higher status families are more apt to have ego enhancing material reinforcements and social reinforcements. Social self-esteem, then, is presumed to be a general evaluation of the self in relation to significant others, and socioeconomic status is but one component of social self-esteem.

A reanalysis of the results of a study by Long, Ziller, and Henderson (in press) involved an equal number of boys and girls of normative age for grade in each Grade 6 to 12 in four schools in Queen Anne's County,

Maryland. The Ss were white and lived in a rural area on the eastern shore of Maryland. The Ss had completed six items of the student form of the SSE. Hollingshead's Occupational Scale based upon head of household's occupation (Hollingshead & Redlick, 1953) was used. Four classifications evolved which provided classes with maximally equivalent numbers of Ss per class. Thus, 83 Ss whose fathers' occupation was professional, business, office worker, or salesman constituted Class 1; 66 Ss whose fathers were classified as skilled labor constituted Class 2; 77 Ss whose fathers were farmers constituted Class 3; and 69 Ss whose fathers were semiskilled or unskilled labor constituted Class 4. The mean self-esteem scores of these Ss were 23.0, 22.4, 20.4, and 20.5, respectively ($F = 2.54$, $df = 3/292$, p about .05). Self-esteem and socioeconomic status are positively associated in this American sample of Ss.

SELF-ESTEEM AND CULTURE

In an analysis of cultural shaping of conception of the self, Hallowell (1955) assumes that the individual's self-image and his interpretations of his own experience cannot be separated from the concept of the self that is normative in his society.

A similar analysis to the foregoing was made possible by the availability of the caste of Indian Ss in reanalysis of a cross-cultural study (Long, Ziller, Ramana, & Reddy, 1966). The Ss consisted of 50 boys and 50 girls from Form I of the M.V.D.M. High School in Vesakapatnam, Andhra, South India. The children ranged in age from 10 to 14 with a median age of 12. Six items of the student form of the SSE were used. The instructions were read aloud in Telugu, the native language of the children, by one of the Indian Es.

The Ss were found to be members of four castes: (a) Caste 1 (Brahmin), $N = 39$, self-esteem = 27.1; (b) Caste 3 (Visya), $N = 9$, self-esteem = 25.1; (c) Caste 4 (included 18 types such as Satani and Najara which are all associated with crafts), $N = 48$, self-esteem = 29.3; (d) Caste 5 (Harijan and Relly), $N = 4$, self-esteem = 30.3. Because of the extremely small number of Ss in two of the castes, Castes 1 and 3 were combined and compared with Castes 4 and 5. The results were

statistically significant at the .05 level of confidence ($F = 5.6$, $df = 1/98$).

The results were in opposition to expectations based upon the results of the previous study. In this sample of Indian Ss the highest self-esteem was expressed by students in the lowest caste (in the results above, the self-esteem scores were simply reversed; high scores represent high self-esteem). However, school was attended within this age group and in the region by only 14% of the population. The children from the lowest castes, then, may be using other children of their caste who do not attend school as points of reference, and within this frame of reference their own status appears extremely high. Thus, as indicated in the initial self-esteem framework, the field of comparison becomes crucial. Self-esteem is defined with regard to significant others.

In this same study of Indian children it was possible to compare the self-esteem of a comparable sample of boys and girls from the Queen Anne's County sample mentioned earlier. Here again the high score represents high self-esteem. The Indian and American means were 28.1 and 20.6 respectively ($F = 65.6$, $df = 1/196$, $p < .005$). Again, however, the privilege of school attendance may be the crucial variable. The Indian student who perceives himself as being a member of a select group by virtue of school attendance may have higher self-esteem than American children, who all attend school as a matter of course. In terms of the discussion of the association between self-esteem and socioeconomic status, it is proffered that the school environment is the most salient status variable for the Indian children.

SELF-ESTEEM AND CONFORMITY

In a summary of the research concerning self-esteem and conformity, Wylie (1961) acknowledges that there is a trend indicating an inverse relationship; individuals with low self-esteem tend to be more persuasible. In the present study 41 high school seniors were administered the six-item student form of the SSE and then were placed in the classic Asch conformity situation (Asch, 1956). Under the nine extreme conditions where the unanimous majority (of four in this study) chose the line

which deviated most from the standard line, the biserial correlation between conformity and the SSE was .32, $p < .05$. Higher self-esteem was associated with higher conformity.

As stated at the outset, previous research was equivocal. Nevertheless, it is compelling to view these results as negative. In view of the strong social component inherent in the present measure of self-esteem, however, it is still possible to interpret the results as indicating that the person with high self-esteem within a social context may not perceive conforming behavior under low cost conditions as damaging to the self system.

SELF-ESTEEM AND THE NEUROTIC PERSONALITY

Ausubel (1952) regards self-esteem as the outcome of achieving a status commensurate with one's conception of self-importance. He proposes that a devaluation of the self-concept is necessary in the face of reality and in order to avoid severe injury to self-esteem. Trauma to self-esteem may result if ego importance is devalued extremely. On the other hand, personality disorders may evolve from untenable notions of omnipotence to which the child is subject.

Results of experiments which have investigated the relationship between adjustment and self-regard (Wylie, 1961) are equivocal. Again, however, the results may simply reflect the shortcomings of verbal self-report measures of self-esteem, particularly when used with neuropsychiatric patients.

In the first and second of three studies involving the SSE (Ziller, Megaw, & DeCencio, 1964), the felt circle approach described earlier was employed. In the first study involving a set of significant other members of an acute neuropsychiatric treatment ward (psychiatrist, psychologist, social worker, nurse, nurses' aide, other patients in the ward, and yourself), patients who had been administered electroconvulsive shock therapy ($N = 10$) from 1 to 6 weeks prior to completing the SSE as compared with those patients who had not received this treatment ($N = 35$) placed the circle representing themselves in the last position to the right more frequently ($\chi^2 = 12.34$, $p < .01$). These results appear to validate the psychiatric screening techniques for

depression since electroconvulsive shock treatment was only recommended for acutely depressed patients.

In the second experiment in the series (Ziller, Megaw, & DeCencio, 1964), 25 patients from the neuropsychiatric ward of an acute treatment center and 23 volunteers from the hospital staff, including personnel from several levels of the occupational hierarchy, arranged in a horizontal straight line on a black felt field 10 significant social elements represented by symbols on a piece of felt 2 inches in diameter. The elements included "mother," "father," "your wife or girl friend," "the most successful person you know," "the happiest person you know," "the person with whom you are most uncomfortable," "employer," "your doctor," and "your friend." The circles were placed in a random order on a table in front of the black felt field. On each circle there was a one-word description of the person whom the circle represented. A list of the social elements that the circles represented was presented in alphabetical order on the same table. The Ss were asked to "place them all on the board in a straight line according to some relationship that you decide upon." Normals were found to have placed the "self" in a higher position than the patients in the assumed left-right hierarchy ($t = 4.57$, $p < .001$).

In the third study in the series (Ziller & Grossman, 1967), 90 male, acute neuropsychiatric patients and 87 male employees of the same hospital served as Ss. Two self-esteem items were administered to the patients during the first week of admission to the hospital. In the first item, S was asked to choose a circle to represent himself from 10 circles arranged in a vertical column. Circles were weighted from 1 to 10, with a higher score associated with a higher position.

The second measure of self-esteem presented a horizontal array of circles. The Ss were also presented with the same list of 10 significant other persons (including the self) used in the preceding study. Here the usual left-right hierarchy was assumed.

The results corroborate earlier findings in the series. Neuropsychiatric patients in comparison with normals show lower self-esteem

on both measures (horizontal, $p < .10$; vertical, $p < .05$).

OVERVIEW

The series of studies described here represents the first phase in a program of research concerning self-other orientation. The program of research involves the integration of a theory of personality involving self-other perceptions and an instrument designed to measure the evaluation of the self in relation to significant others using topological representations of self and others. The outcome of the present approach, although balanced with regard to the emphasis on theory, instruments, and research, rests largely upon the utility of the measures involved. Measurement remains the missing link in personality research. With regard to self-social constructs, measurement is dependent upon the method of communicating between *S* and the scientist. Here we have proposed that there are some distinct advantages to avoiding the usual verbal self-report approach and substituting for it a topological approach with limited verbal demands.

The results tend to support the validity and utility of the approach to the measure of self-esteem. Social objects with greater value tend to be placed to the left in the horizontal display; the absolute difference between location of self and a low-status social object is significantly associated with the left to right location of the self; left-right location of the self is significantly associated with the up-down location of the self. Higher self-esteem was found to be associated with social acceptance, social participation, socioeconomic status (only in an American sample), identification with parents, consistency of social behavior, and the normal as opposed to the neurotic personality. Finally, winning political candidates for state legislative offices rose in self-esteem, whereas losing candidates dropped in self-esteem.

The results of three studies require reconsideration. In one of these negative or very low correlations were found between the SSE and three frequently used measures of self-esteem. The results are consistent with those of an earlier study, however, and strongly suggest that the SSE measures an aspect of

self-evaluation which is in a different factor region.

The second result requiring reconsideration concerns the tendency of low caste as compared with high caste Asian Indian students to place the symbol representing the self more to the left. By way of explanation, it was observed that only a select few of the low caste as opposed to a high percentage of the high caste attended public schools, and as a consequence, the high status of the selected low caste student in comparison with the non-selected low caste student was reflected in their self-evaluations on the SSE. Finally, persons with high social self-esteem were found to be *more* conforming in the Asch situation. It is difficult to discount this finding, even though previous studies concerning a relationship between self-esteem and conformity have been equivocal. Against the background of the directly supporting results, the three results requiring qualification do not appear to require a reexamination of the validity of the concept of social self-esteem.

The social context of self-esteem has been emphasized in the present approach, and, indeed, the results of most of the studies are concerned with social behavior (i.e., popularity, frequency of participation in group discussions, parent-child relationships, conformity). An attempt was made to describe the concept used here in terms of its social correlates. Thus, the term "social self-esteem," was used. The meaning of social self-esteem as it evolves against the background of its social correlates suggests social acceptance or perhaps self-other confidence. The individual who is assured of his high self-evaluation within a social context is more consistent in social behavior and more accepted by others. As stated at the outset, self-acceptance and social acceptance are intrinsically interdependent.

If, then, the traditional self-report measure of self-esteem is unrelated to social self-esteem, a reanalysis of the meaning of the traditional measures is indicated. It is now suggested that the self-report measures indicate a socially desirable self-esteem, an evaluation of the self that the reporter is willing to reveal, or that he desires the other to accept.

Aside from the question of the social con-

text of self-esteem, the limited verbal demands of the present approach recommend it as a most useful measure in cross-cultural research, developmental research, and research in general where there may be some question of the comparable verbal ability of the Ss in relation to the experimental tasks.

The fundamental assumption of the proposed measure of self-esteem is the proclivity, in the Western culture, for left to right linear ordering of objects. The findings of the present series of studies along with a long history of research appears to support this assumption. As has already been noted, the tendency to attribute greater importance to the object placed at the extreme left position in a horizontal display was recorded by Morgan in 1944. Introspectionists such as Lashley (1961) and Inhelder and Piaget (1958, p. 252) have also noted the use of spatial imagery in thinking about nonspatial orderings.

The most systematic work in this area has been conducted by DeSoto (1960, 1961; DeSoto, London, & Handel, 1965). For example, in spatial imagery associated with syllogisms, the results suggest that the left end is the preferred starting point for a horizontal ordering (DeSoto, London, & Handel, 1965).

The studies presented here have attempted to extend this work to the development of a universal communication system; a sign language, if you will, for describing self-other orientations. On this assumption, a number of self-other configurations have been developed (see Ziller & Grossman, 1967; Ziller & Long, 1966) using other spatial arrangements of symbols representing self and significant others which are designed to measure self-other power orientation, marginality, social interest, identification, identification with the majority, self-centeredness, social inclusion, and openness.

REFERENCES

- ASCH, S. E. Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs*, 1956, 70(9, Whole No. 416).
- AUSUBEL, D. P. *Ego development and the personality disorders*. New York: Grune & Stratton, 1952.
- BILLS, R. E., VANCE, E. L., & MCLEAN, O. S. An index of adjustment and values. *Journal of Consulting Psychology*, 1951, 15, 257-261.
- BORGATTA, E. F. A systematic study of interaction process scores, peer and self assessment, personality and other variables. *Genetic Psychological Monographs*, 1962, 65, 219-291.
- COOMBS, C. H., RAIFFA, H., & THRALL, R. M. Some views on mathematical models and measurement theory. *Psychological Review*, 1954, 61, 132-144.
- COOPERSMITH, S. A method for determining types of self-esteem. *Journal of Abnormal and Social Psychology*, 1959, 59, 187-194.
- COOPERSMITH, S. *The antecedents of self-esteem*. San Francisco: Freeman, 1967.
- CROWNE, D. P., & MARLOWE, D. *The approval motive*. New York: Wiley, 1964.
- CROWNE, D. P., STEPHENS, M. W., & KELLY, R. The validity and equivalence of tests of self acceptance. *Journal of Psychology*, 1961, 51, 101-112.
- CUTICK, R. A. Self-evaluation of capacities as a function of self-esteem and the characteristics of a model. Unpublished doctoral dissertation, University of Pennsylvania, 1962.
- DESOTO, C. B. Learning a social structure. *Journal of Abnormal and Social Psychology*, 1960, 60, 417-421.
- DESOTO, C. B. The predilection for single ordering. *Journal of Abnormal and Social Psychology*, 1961, 62, 16-23.
- DESOTO, C. B., & KUETHE, J. L. Subjective probabilities of interpersonal relationships. *Journal of Abnormal and Social Psychology*, 1959, 59, 290-294.
- DESOTO, C. B., LONDON, M., & HANDEL, S. Social reasoning and spatial paralogic. *Journal of Personality and Social Psychology*, 1965, 2, 513-521.
- DIGGORY-FARNHAM, S. Self-evaluation and subjective life expectancy among suicidal and nonsuicidal psychotic males. *Journal of Abnormal and Social Psychology*, 1964, 69, 628-634.
- FAIRWEATHER, G. W. *Social psychology in treating mental illness*. New York: Wiley, 1964.
- FESTINGER, L. A theory of social comparison processes. *Human Relations*, 1954, 7, 117-140.
- HALLOWELL, A. I. *Culture and experience*. Philadelphia: University of Pennsylvania Press, 1955.
- HENDERSON, E. H., LONG, B. H., & ZILLER, R. C. Self-social constructs of achieving and non-achieving readers. *The Reading Teacher*, 1965, 19, 114-118.
- HOLLINGSHEAD, A. B., & REDLICK, F. C. Social stratification and psychotic disorders. *American Sociological Review*, 1953, 18, 163-169.
- INHELDER, B., & PIAGET, J. *The growth of logical thinking*. New York: Basic Books, 1958.
- KELLY, G. A. *The psychology of personal constructs*. New York: Norton, 1955.
- KUETHE, J. L. Social schemas. *Journal of Abnormal and Social Psychology*, 1962, 64, 31-36.
- LASHLEY, K. S. The problem of serial order in behavior. In S. Saporta (Ed.), *Psycholinguistics*. New York: Holt, Rinehart & Winston, 1961.
- LEWIN, K. *A dynamic theory of personality*. New York: McGraw, 1935.
- LONG, B. H., ZILLER, R. C., & HENDERSON, E. H. Developmental changes in the self-concept during middle childhood. *School Review* (in press).

- LONG, B. H., ZILLER, R. C., RAMANA, K. V., & REDDY, J. E. Self-social orientations of Indian and American children. Unpublished manuscript, Goucher College, 1966.
- MANN, R. D. A review of the relationship between personality and performance in small groups. *Psychological Bulletin*, 1959, 56, 241-270.
- MORGAN, J. J. B. Effects of non-rational factors in inductive reasoning. *Journal of Experimental Psychology*, 1944, 34, 159-168.
- MOSSMAN, B. M., & ZILLER, R. C. Self-esteem and consistency of social behavior. *Journal of Abnormal Psychology*, 1968, 73, 363-367.
- PIAGET, J. *The psychology of intelligence*. (Trans. by M. Piercy & D. E. Berlyne) London: Routledge & Kegan Paul, 1947.
- RIDGEWAY, S. The relationship among three measures of self concept. Unpublished master's thesis, University of Maryland, 1965.
- ROGERS, C. R. *Client-centered therapy*. Boston: Houghton Mifflin, 1951.
- ROSENBERG, M. *Society and the adolescent self-image*. Princeton: Princeton University Press, 1965.
- WITKIN, H. A., DYK, R. B., FOTERSON, H. F., GOODENOUGH, D. R., & KARP, S. A. *Psychological differentiation*. New York: Wiley, 1962.
- WYLIE, R. C. *The self concept*. Lincoln: University of Nebraska Press, 1961.
- ZILLER, R. C., ALEXANDER, M., & LONG, B. H. Self-social constructs and social desirability. Unpublished manuscript, University of Delaware, 1964.
- ZILLER, R. C., & GROSSMAN, S. A. A developmental study of the self-social constructs of normals and the neurotic personality. *Journal of Clinical Psychology*, 1967, 23, 15-21.
- ZILLER, R. C., & LONG, B. H. Self-social constructs: Theory and instruments. Unpublished manuscript, University of Oregon, 1966.
- ZILLER, R. C., MEGAS, J., & DECENCIO, O. Self-social constructs of normals and acute neuropsychiatric patients. *Journal of Consulting Psychology*, 1964, 28, 50-63.

(Received March 8, 1968)

(Continued from page 77)

- Comparative Study of the Wechsler Preschool and Primary Scale of Intelligence and the Stanford-Binet Intelligence Scale, Form L-M, among Culturally Deprived Children: Allan Barclay* and Allan C. Yater: Cardinal Gennon Memorial Hospital for Children, 1465 South Grand Boulevard, St. Louis, Missouri 63104.
- Measurement of Social-Evaluative Anxiety: David Watson* and Ronald Friend: Psychology Department, University of Hawaii, Honolulu, Hawaii 96822.
- Effect of Physical and Cognitive Somatic Arousal on Rorschach Responses: An Experimental Test of the Assumption that Body Image Influences the Perceptual Organization of Unstructured Stimuli: Steven N. Van De Mark and Charles Neuringer*: Department of Psychology, University of Kansas, Lawrence, Kansas 66044.
- Object-Discrimination Learning Set Acquisition in Young Cerebral Palsied Children in Relation to Tested and Rated Intelligence: Douglas Pearson*: California State College at Long Beach, 6101 E. Seventh Street, Long Beach, California 90801.
- Conditioning and Transfer of Affective Self References in a Role Played Counseling Interview: Robert J. Hoffnang*: Counseling Service, University of Cincinnati, Cincinnati, Ohio 45221.
- Revised Suicide Potential Scale: R. W. Miskimins and Lowell T. Wilson*: Mental Health and Manpower Project, 3520 West Oxford Avenue, Denver, Colorado 80236.
- Different Psychological Effects of Lateralized Brain Damage: Oscar A. Parsons*, Arthur Vega, Jr., and Julian Burn: The University of Oklahoma Medical Center, 800 North-east Thirteenth Street, Oklahoma City, Oklahoma 73104.
- Predicting the Success of Peace Corps Community Development Workers: Charles Dicken*: Department of Psychology, San Diego State College, San Diego, California 92115.

* Asterisk indicates author for whom address is supplied.

NEED FOR APPROVAL AND THREATENED NEGATIVE EVALUATION AS DETERMINANTS OF EXPRESSIVENESS IN A PROJECTIVE TEST

HERBERT M. LEFCOURT¹

University of Waterloo, Ontario, Canada

Constriction, or the lack of expressiveness in projective testing as measured by the number of human movement responses given to the Barron inkblot test, was predicted on the basis of Ss' need for approval and expectancy of possible disapproval which derived from E's directions. Results strongly supported the emphasis upon situation cues and personal motives as predictors of responses to projective testing. Additional findings with repression-sensitization and the California Psychological Inventory-Good Impression scale complimented the findings with need for approval. Constriction was most evident when tests offered personal threat but this was especially so for high need for approval persons seeking to make good impressions.

Constriction, a lack of productivity, and stereotyped responding to projective tests have usually been identified as avoidant, defensive ways of coping with evaluation situations. Defensiveness and constriction on such tests as the Rorschach commonly are said to indicate attempts on the examinees' part to avoid criticism and to protect their self-images. From such interpretations it might be inferred that examinees would behave differently in projective testing if they did or did not believe that the potential for criticism was implicit in the testing procedure. Indeed, there has been considerable research into the effects of situational cues, which could alter such beliefs, upon projective testing. Masling (1960) and Rotter (1960) have each reviewed the research literature in this area and report substantial support for the hypothesized context effects upon projective test performance. Rotter (1963) has stated that S "attributes some purpose to the test, and the purpose he attributes to it may well hinge on odd bits of information, the setting of the test, differential emphasis on the words in the same instructions, and so forth [p. 805]." In other words, the ambiguity in projective testing produces a greater search for cues by Ss; therefore, Ss become more easily influenced by situational cues than they would be if the tests purposes were more explicit.

Despite this well-supported contention concerning the effect of situational cues upon projective testing, there has been little research reported which might further illuminate this finding. For instance, the questions could be asked: Who, or what sort of person would respond to what sorts of cues in projective testing? In essence, when we receive a constricted and defensive production from a patient on an inkblot test, can we reliably attribute his defensiveness to a "personality style," to threats implicit in the evaluation outcomes for him, or to some combination of these? As of the present, most clinicians would probably assume that defensiveness is a trait exhibited by persons cross-situationally. Broad categorizations of defensiveness such as repression-sensitization are developed with the assumption that persons behave similarly across different situations. When subjected to experimental variations, however, even predictions with this widely employed personality variable have been found to vary, given different cues in criterion situations such as projective testing (Lefcourt, 1966).

The present investigation was designed to test out one theoretical approach to the prediction of defensiveness in projective testing. With variables derived from social learning theory (Rotter, 1954), we will attempt to predict the degree of defensiveness in an inkblot test on the basis of an S's needs and expectancies as to how his test behavior may reflect back upon himself.

¹The author wishes to express appreciation to Gary Eason, who helped at different stages of the analysis.

In three recent investigations concerning productivity upon inkblot tests, the production of human movement (M) responses has been found to be a useful indicator of expressiveness (the opposite of defensiveness). In a study by Lefcourt and Steffy (1966) human movement responses were given more often to an examiner from whom there appeared to be less likelihood of censure. Since cognitive style investigators have reported that M responses are positively correlated with active visual scanning (Luborsky, Blinder, & Schimek, 1965), negatively correlated with defensiveness in the report of conflict stimuli (Kagan & Moss, 1962), and negatively related to repressive or denial type defenses (Gardner, Holzman, Klein, Linton, & Spence, 1959; Levine & Spivack, 1960), these writers assumed that M responses reflect greater participation, effort, and consequently, expressiveness or the opposite of defensiveness by Ss performing on inkblot tests. That more M responses were obtained from Ss who were less likely to anticipate censure gave some support to these hypotheses.

In a microanalysis of M response production, Megargee (1966) found that the number of words used to define a percept is one of the factors necessary for producing M responses. While he did not interpret his data in this fashion, Megargee's findings suggest that M production requires greater verbal expressiveness, the "willingness to talk or write" which, in turn, indicates a lack of perceived threat in the examination situation.

A third study of immediate relevance is an unpublished dissertation by Tutko (1962) reported by Crowne and Marlowe (1964). In this study, psychotic patients were subjected to a battery of projective tests under threatening versus supportive conditions. In the former condition the tests were described as measures of "mental sickness." The supportive directions emphasized a concern with the tests per se assuring anonymity to the patient. Prior to testing the patients had completed the Marlowe-Crowne (M-C) scale of need for approval which is a Lie scale type of measure in which Ss can endorse low probability but highly desirable self-descriptive comments. A person scoring high on this measure is said to need or value the approval

of the examiner. The patients were then tested on the Rorschach, TAT, and Incomplete Sentences tests. Responses on these tests were judged for "revealingness," "pathology," and "defensiveness." While the results vary by test and rating, generally the threat condition produced less revealingness and pathology, and greater defensiveness, especially among patients who had scored high in the need for approval. In regard to M responses which were not subjected to the detailed analyses as were the ratings, high need for approval patients produced less M than those who scored low on the M-C scale.

From these previous studies, then, we may hypothesize that when Ss are confronted with inkblot tests, they can be readily influenced by E's directions as to the test's purposes. Since M responses reflect the degree of expressiveness engaged in by S, then this type of response should vary as a function of E's directions, especially when those directions suggest that negative evaluations can be derived from S's expressions. This hypothesis is not novel and simply reiterates previous findings as to the effects of threat upon Ss engaged in projective testing. In addition, we are hypothesizing that Ss who have expressed a need for approval from the examiner, as indicated by their performance on the M-C scale, will exhibit considerably more constriction and defensiveness in the face of threatened negative evaluation than those who have not expressed such a need for approval. Specifically, we are predicting that constriction and lack of productivity, as measured by the amount of M responses given to an inkblot test, will be a function both of S's needs (approval) and the specific expectancy derived from directions that his behavior in the testing situation may cause him grief in regard to the fulfillment of that need.

A supplementary hypothesis was added late in the investigation. Boor and Schill (1967) report that Ss who describe themselves as low in anxiety but who are high in the need for approval exhibit considerably more task-anxiety than their low need for approval peers. Since scores on Byrne's (1961) Repression-Sensitization (R-S) scale were available for all Ss, supplementary hypotheses were constructed concerning the interaction

of the M-C and R-S scales upon M responses under different conditions. It was predicted that low-anxious (repression)-defensive (high M-C) Ss would be more responsive to threat directions than other Ss, that they should produce less M than any other of the three groups (low M-C-repressor, low M-C-sensitizer, high M-C-sensitizer) when anticipating negative evaluation possibilities.

METHOD

Subjects

The sample for this investigation was 133 Ss, 71 males and 62 females, who were members of the author's third-year undergraduate-level course in personality at the University of Waterloo.

Procedure

During the semester, the entire class completed several test measures which were used to illustrate given theoretical orientations. Among these measures was the M-C scale of need for approval and Byrne's R-S scale. The overall class mean for the former was 14.97, $SD=5.46$, which is comparable to Crowne and Marlowe's (1964) norms. The overall mean was 56.76, $SD=15.39$ for the R-S variable. In later class sessions the group was divided in half alphabetically. One group was informed that they were to take an inkblot test which was not the Rorschach. On this test, they were instructed to be as imaginative as possible. They were told that there would be 26 inkblots shown for a minute apiece. Barron's M inkblot test (Barron, 1955) was then presented by slide projector. The second group was told that they were to take an inkblot test which was not the Rorschach, but which had derived from that test. It was explained that while the inkblot type tests had originally been devised to measure such attributes as mental illness, this should not really be relevant to the present group. The students were then instructed to be as imaginative as they could be. Again Barron's M measure was presented by slide projector.

The specific predictions were that Ss who were high in the need for approval would give less M responses than would Ss who were low in the need for approval, especially when a threat of negative evaluation was present. Since the label of mental illness is commonly associated with disapprobation, these directions would seem to be particularly arousing to those Ss with a high need for approval.

Secondarily, it was predicted that among the high need for approval Ss, those who described themselves as low in anxiety (repressors) would produce even less M responses than their high need for approval peers who admit to anxiety. While the R-S scale employed here is not identical to the Taylor Manifest Anxiety scale used in the Boor and Schill (1967) article, it has been found to be strongly correlated ($r=.91$, Joy, 1963) with that scale.

The first and more benign condition described above will be referred to as the creative imagination (CI) condition and the threat session as the mental illness (MI) condition.

RESULTS

The number of M responses were scored by two raters who established a high rate of agreement ($r=.92$, $p<.01$) using the criteria developed by Spivack, Levine, and Sprigle (1958). With a median split of scores on the M-C scales ($Mdn=15$) Ss were classified as low if their scores fell in the 0 to 14 range, and high in the need for approval if their scores were 15 and above. Sixty-three Ss were classified as low, 70 as high in the need for approval; 73 Ss were run in the CI and 60 in the MI conditions. The analyses of variance reported, therefore, are based on an unweighted mean method for unequal N analyses (Winer, 1962), and the design is that of a 2×2 factorial including two levels of need for approval and two conditions (CI vs. MI).

Two dependent measures are derived from Barron's inkblot test: the number of M responses and the M threshold score. The latter measure is defined by the number of the card when S produces his first M. Since the cards are ordered for their pull for M, a high threshold score would indicate that S produces his first M response late in the series and can be said to require highly explicit stimuli to evoke an M. A low threshold would indicate a readiness to produce M responses early in the series when only minimal cues suggestive of M are present. While the two measures, number of and threshold for M must of necessity be related, there is some evidence to suggest that they are different enough to require separate attention (Nyman, 1961; Spivack, Levine, & Sprigle, 1958).

Table 1 presents the means and standard deviations of the number of M responses and of M threshold scores.

As indicated in Table 2, there is a highly significant main effect. The Ss who are high give less M than those who are low in the need for approval. This effect is largely accounted for by the very evident constriction of high need for approval Ss in the MI condition as is indicated by the significant inter-

TABLE 1

MEANS AND STANDARD DEVIATIONS OF TOTAL M RESPONSES AND M THRESHOLD IN CI AND MI CONDITIONS BY Ss LOW AND HIGH IN THE NEED FOR APPROVAL

Condition	Low need for Approval	High need for Approval
Total M		
CI		
<i>M</i>	11.31	12.21
<i>SD</i>	2.35	3.38
MI		
<i>M</i>	11.14	8.94
<i>SD</i>	2.97	2.47
Threshold for M		
CI		
<i>M</i>	6.97	8.14
<i>SD</i>	3.35	4.51
MI		
<i>M</i>	8.73	10.28
<i>SD</i>	4.95	5.50

Note.—For low need for approval, $N = 63$; for high need for approval, $N = 70$. CI = creative imagination; MI = mental illness.

action term. Subsequent simple effects analyses reveal that the interaction largely derives from differences between high need for approval Ss across the two conditions ($p < .01$), and between high and low need for approval groups in the MI condition ($p < .001$). It is evident that the hypotheses regarding need for approval receive strong support from the first analysis.

In the analysis of variance for M threshold presented in Table 3, the pattern of results appears similar to that for the number of M responses. The main effect for need for approval again reaches statistical significance. The Ss with a high need for approval emit

their first M later than Ss low in the need for approval. A trend is also found indicating that threshold scores tend to be higher in the MI condition for both groups while no interaction is obtained between condition and need for approval. In general, the results obtained are less impressive with the threshold measure. The large variance for this as opposed to the number of M responses measure may help to explain the lesser power of threshold as a dependent variable.

In the supplementary analysis which includes the R-S variable, low-anxious Ss or repressors were defined as those whose R-S score was below the median (0-55), sensitizers above (56 and above). The means, standard deviations, and N s for this three-factor design are presented in Table 4. The analysis conducted was a $2 \times 2 \times 2$ unweighted mean analysis for unequal N s. While the results obtained for the need for approval variable were the same as that reported above, the predicted interaction between need for approval, R-S, and CI versus MI was obtained ($F = 6.59$, $df = 1/123$, $p < .05$). In the subsequent simple effects analyses, no differences were obtained among groups in the CI condition. In the MI condition, however, significant differences were obtained ($p < .01$). To begin with, the low need for approval-repressor Ss produced more M responses than all of the other groups while the high need for approval-repressor produced less M than the other groups. These differences were all significant at least at the .05 level except for that between the high need for approval-

TABLE 2

ANALYSIS OF VARIANCE OF EFFECTS OF NEED FOR APPROVAL AND CI VERSUS MI CONDITIONS UPON THE NUMBER OF M RESPONSES PRODUCED

Source	<i>df</i>	<i>MS</i>	<i>F</i>
Need for approval (A)	1	83.05	9.68*
CI vs. MI conditions (B)	1	15.97	1.86
A \times B	1	66.20	7.72*
Error	129	8.58	

Note.—CI = creative imagination; MI = mental illness.

* $p < .01$.

TABLE 3

ANALYSIS OF VARIANCE OF EFFECTS OF NEED FOR APPROVAL AND CI VERSUS MI CONDITIONS UPON M THRESHOLD SCORES

Source	<i>df</i>	<i>MS</i>	<i>F</i>
Need for approval (A)	1	125.7106	5.57*
CI vs. MI conditions (B)	1	64.1926	2.84**
A \times B	1	1.1935	1 <
Error	129	22.5660	

Note.—CI = creative imagination; MI = mental illness.

* $p < .05$.

** $p < .10$.

TABLE 4

MEANS AND STANDARD DEVIATIONS OF TOTAL HUMAN MOVEMENT RESPONSES IN CI AND MI CONDITIONS BY Ss CLASSIFIED BY REPRESSION-SENSITIZATION AND NEED FOR APPROVAL

Ss	CI	MI
Low need for approval		
Repressors		
<i>M</i>	10.73	13.05
<i>SD</i>	1.60	3.61
<i>N</i>	11	19
Sensitizers		
<i>M</i>	11.67	10.81
<i>SD</i>	2.64	2.90
<i>N</i>	18	16
High need for approval		
Repressors		
<i>M</i>	11.29	8.37
<i>SD</i>	3.59	2.32
<i>N</i>	17	19
Sensitizers		
<i>M</i>	10.83	10.00
<i>SD</i>	2.34	2.32
<i>N</i>	18	13

Note.—CI = creative imagination; MI = mental illness.

repressors and sensitizers. Nevertheless, this difference was in the predicted direction. In short, the repressors or low anxious persons were both the lowest and highest in *M* production with need for approval being the variable that predicts the direction. When the groups are compared across conditions, the low need for approval-repressors are found to increase their *M* responding in the MI condition ($p < .05$) while the high need for approval-repressors decrease in the number of *M* given in the MI as compared to the CI condition ($p < .01$).

That the conditions play such an important role in the interaction with the personality variables in predicting *M* responses receives further corroboration from results with the California Psychological Inventory (CPI; Gough, 1964) which had been administered to Ss earlier in the year. The variable Good Impression (GI) on the CPI would seem to be similar to the need for approval. The Ss scoring high on GI are said to be concerned about how others react to them and wish to make a good impression. With the total sample of 129, the M-C need for approval scale was found to be related to GI, $r = .41$,

$p < .01$. In the CI condition, GI and the number of *M* responses fall short of significance ($r = -.19$, *ns*). However, in the MI condition, GI and *M* are significantly related ($r = -.34$, $p < .01$). In other words, when it becomes clear that there is a possibility of negative evaluation, the wish to create a good impression becomes more salient as a predictor to the manner in which one will be expressive.

DISCUSSION

The results of the present investigation give strong support to the original contention that small cues present in a projective testing situation can dramatically affect Ss' performances. That we have been able to predict the interaction between a personality variable such as need for approval and the response to a particular kind of situational cue (threat of negative self-evaluation) suggests that more exacting methods for interpreting test behavior may be possible. Rather than relying on a trait assumption in interpreting test results (e.g., he was defensive on the test so that he is henceforth described as being defensive), it may become possible to interpret test behavior as a function of the situational cues in testing which can be controlled in interaction with personality variables which are more articulated in given personality theories and supported with empirical research.

The results of this study indicate that the absence of *M* responses can be viewed as indicators of constriction and defensiveness. We are not arguing that *M* is the only or even the best determinant for inferring defensiveness, but that it has proven useful as a measure of defensiveness in this case as well as in previous studies (Lefcourt & Steffy, 1966). *M* responses apparently reflect the kind of participation and expressiveness which Ss might fear would reveal them more and consequently make them more vulnerable to an examiner's judgment. This is not to say that Ss are conscious of a relationship between *M* and self-disclosure. Rather, it is assumed that *M* responses require a somewhat greater effort and participation than a simple form response, and that this effort will not be expended when Ss are on the defensive.

The data obtained with the GI measure from the CPI lends support to that obtained

with the need for approval scale, that the incidence of M responses varies as a function of Ss' intentions. The Ss who desire to create a good impression are less expressive, or more constricted, when expression is perceived as constituting a risk to the impression they wish to maintain.

It is interesting that the M threshold measure does not appear to generate the sorts of results obtainable with the total number of M responses. Barron (1955) originally devised his test to create this threshold measure. Other investigators such as Spivack et al. (1958) have noted that very few Ss produce an M on every card following the threshold card. Consequently, there tends to be a less than perfect congruence between the two M measures. Nyman (1961) also found that number of M and M threshold were related in different ways to a series of measures related to the tolerance of delay. In the present study, M threshold is found to be a less sensitive measure than the total M score. The data, though in the same directions as the number of M, do not reach statistical significance possibly because of the larger variance secured with the threshold measure.

The additional findings obtained with the repression-sensitization measure lead to interesting hypotheses. Since the present result supports the findings of Boor and Schill (1967) in singling out the high need for approval-low anxious Ss (repressors) as a deviant group, further conjecturing seems justifiable. It is possible to ask why persons who report that they are unanxious, or less uncertain about who and what they are, should at the same time be desirous of external reassurance. A possible answer is that these persons are describing themselves as nonanxious because that too is seen as necessary for the securing of approval. In other words, such persons would find the admission of anxiety as a threat to their maintenance of social approval. In effect, the very feelings of uncertainty or anxiety might prove, ironically, to be further anxiety arousing for such persons. Consequently, when these persons are being evaluated on a dimension relevant to anxiety, their first predisposition would be to avoid self-disclosure as would be evidenced in constriction and withdrawal. The fact that

the low need for approval-repressor group, on the other hand, significantly increases their tendency to give M responses with threatened negative evaluation may reflect the fact that this group feels they have nothing to hide. If they do not seek approval and yet deny anxiety, conceivably these are less troubled individuals who have confidence in their own well-being. Consequently, when being evaluated their predisposition is to be as open or self-disclosing as possible. In other words, if viewed enough, they will in all probability be judged favorably. However, it might also be argued that this group simply does not care what the evaluator believes since they are low in the need for approval. Nevertheless, the unpredicted increase in M responding with the anticipation of negative evaluation which this group exhibits would tend to belie this assertion.

In conclusion, we have obtained significant support for our original hypotheses. Situational cues, as slight as they were in the present study, have a large but differential impact upon Ss undergoing projective testing. The results likewise reflect the usefulness of social learning theory in predicting the kinds of behavior that often come to the clinician's attention.

REFERENCES

- BARRON, F. Threshold for the perception of human movement in inkblots. *Journal of Consulting Psychology*, 1955, 19, 33-38.
- BOOR, M., & SCHILL, T. Digit symbol performance of subjects varying in anxiety and defensiveness. *Journal of Consulting Psychology*, 1967, 31, 600-603.
- BYRNE, D. The Repression-Sensitization Scale: Rationale, reliability, and validity. *Journal of Personality*, 1961, 29, 334-349.
- CROWNE, D. P., & MARLOWE, D. *The approval motive: Studies in evaluative dependence*. New York: Wiley, 1964.
- GARDNER, R., HOLZMAN, P. S., KLEIN, G. S., LINTON, H., & SPENCE, D. P. Cognitive control. *Psychological Issues*, 1959, 1(Whole No. 4).
- GOUGH, H. G. *California Psychological Inventory manual*. Palo Alto: Consulting Psychologists Press, 1964.
- JOY, V. L. Repression-sensitization and interpersonal behavior. Paper read at the meeting of the American Psychological Association, Philadelphia, September 1963.
- KAGAN, J., & MOSS, H. A. *Birth to maturity*. New York: Wiley, 1962.

- LEFCOURT, H. M. Repression-sensitization: A measure of the evaluation of emotional expression. *Journal of Consulting Psychology*, 1966, 30, 444-449.
- LEFCOURT, H. M., & STEFFEY, R. A. Sex linked censure expectancies in process and reactive schizophrenics. *Journal of Personality*, 1966, 34, 366-380.
- LEVINE, M., & SPIVACK, G. *The Rorschach index of repression*. Devon, Pa.: Devereux Foundation, 1960.
- LUBORSKY, L., BLINDER, B., & SCHIMEK, V. Looking, recalling, and GSR. *Journal of Abnormal Psychology*, 1965, 70, 270-281.
- MASLING, J. The influence of situational and interpersonal variables in projective testing. *Psychological Bulletin*, 1960, 47, 65-85.
- MEGARGEE, E. I. The relation of response length to the Holtzman inkblot technique. *Journal of Consulting Psychology*, 1966, 30, 415-419.
- NYMAN, L. The relationship between the ability to perceive human movement response in the Barron inkblot test and the reaction to delay in word-naming and reaction time situations. Unpublished doctoral dissertation, New York University, 1961.
- ROTTER, J. B. *Social learning and clinical psychology*. New York: Prentice-Hall, 1954.
- ROTTER, J. B. Some implications of a social learning theory for the prediction of a goal directed behavior from testing procedures. *Psychological Review*, 1960, 67, 301-316.
- ROTTER, J. B. A historical and theoretical analysis of some broad trends in clinical psychology. In S. Koch (Ed.), *Psychology: A study of a science*. Vol. V. New York: McGraw-Hill, 1963.
- SPIVACK, G., LEVINE, M., & SPRIGLE, H. Barron M threshold values in emotionally disturbed adolescents. *Journal of Projective Techniques*, 1958, 22, 446-449.
- TUTKO, T. A. Need for social approval and its effect on responses to projective tests. Unpublished doctoral dissertation, Northwestern University, 1962.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

(Received March 13, 1968)

CONCEPT OF SELF AND POSTHOSPITAL VOCATIONAL ADJUSTMENT¹

K. L. BERRY

AND

R. W. MISKIMINS²

University of Colorado

Colorado State University

The purpose of this research was to demonstrate the predictive efficiency of self-concept measurement for the posthospital vocational adjustment of psychiatric patients. A total of 104 Ss—vocationally unplaceables, failures, successes, and normals—were administered the Miskimins Self-Goal-Other Discrepancy Scale. There was a clear difference as to psychopathology, i.e., discrepancies among the three aspects of self measured by the MSGO, between normals and psychiatric patients. Unplaceable patients were shown to have high Self-Other minus discrepancies, indicating, generally, psychotic symptomatology, and specifically, difficulties in relating to other persons. The failure group's notable distinguishing characteristic was high Self-Goal minus discrepancy, this suggesting extensive use of manipulative behavior, while successes were most easily identified by high Self-Goal plus discrepancy, this related most often to psychoneurotic diagnoses.

Many forms of clinical data have been utilized for the prediction of posthospital psychiatric adjustment; however, none of these variables has been impressively effective. The present study will attempt to demonstrate that a person's concept of self is a highly effective indicator of one key aspect of psychiatric rehabilitation—posthospital vocational adjustment.

There are various general theoretical formulations relating the "self-concept" to human behavior, psychotherapy, and psychological research; however, if this particular type of theorizing is to prove genuinely fruitful, it must deal in detail with psychopathology. That is, self-concept theory should offer a framework in which all the various deviations from normal behavior may be clearly understood. A review of the literature will show that traditional phenomenal self-concept variables allow only limited differentiation between normals and the major psychodiagnostic groups, neurotic, psychotic, and character disorder (e.g., Miskimins & Simmons, 1966; Rogers & Dymond, 1954; Sarbin & Rosenberg, 1955), and among these three

diagnostic groups (e.g., Friedman, 1955; Hillson & Worchel, 1957). Ibelle (1960) has demonstrated that it is not possible, using typical-self-concept measurement, to distinguish most forms of schizophrenia from normals, and the same problem arises with differentiation of normals and character disorders. Miskimins (1967), however, introduced a novel measurement technique, the Miskimins Self-Goal-Other Discrepancy Scale (MSGO) which, in addition to allowing duplication of variables typically measured, permits the measurement of other self-concept variables which serve to discriminate heretofore undistinguishable psychopathological categories. For example, it provides a self-goal factor which duplicates the typical self-ideal *Q* sort and correlates highly with neurotic psychopathology, and a self and other's evaluation of self factor which clearly distinguishes schizophrenia.

The MSGO is a technique for measuring the discrepancies in a person's alignment of: (a) His self-concept, (b) his goal self-concept, and (c) his perception of how others evaluate him, on a given area of concern. Miskimins (1967) has postulated that the measurement of the incongruencies among the three self-constellations will reflect internal tension, that is, anxiety and the individual's general level of adjustment. Specifically, there are available two kinds of discrepancy scores,

¹This research was supported in part by the United States Department of Labor (OMPER Grant Number 82-06-66-62).

²The authors wish to express their appreciation to Alice Austin, Jane London, Betty Lusky, and Andrea Mattice for their assistance in data analysis and manuscript preparation.

self-goal and self-other. When an individual's aspired goals are those typical to the society (e.g., Friendly and Warm) a plus discrepancy obtains, whereas culturally atypical goal choice (e.g., Unfriendly and Cold) results in a self-goal minus discrepancy. Self-other discrepancy is also bidirectional; that is, when others are seen as viewing one's self very favorably this produces plus discrepancy, as opposed to the situation where others are perceived as hypercritical which produces minus discrepancy. The above four kinds of scores—Self-Goal plus, Self-Goal minus, Self-Other plus, Self-Other minus—represent the first four factors of the MSGO (see Table 1 for a summary of the MSGO factors). Factor 5 of the MSGO is similar to Factor 4 in that it is based upon self-other minus discrepancy, but unique in that it involves only very general, global self-related constructs (e.g., Successful in Life). The last factor considered in this paper is that labeled "Negativism" and operationally denotes those Ss who claim to desire to be more tense, less self-confident, and the like. For further information regarding the MSGO, see Miskimins (1967, 1968).

For the purpose of this study vocational

adjustment has been operationalized in terms of "placeability" and "placement durability," that is, a patient has or has not obtained a job, and he has kept or lost an obtained job. The specific predictions hypothesized for the relationship of the self-concepts of psychiatric patients and their future vocational adjustments are: (a) The placeable group will exhibit significantly higher discrepancy on the Self-Goal plus and minus factors, and significantly lower discrepancy on the remaining four MSGO factors (all Self-Other discrepancies and Negativism) than will unplaceable patients. (b) Among those placed, the vocational success group will manifest significantly higher discrepancy on the Self-Goal plus and Self-Other plus factors and will score significantly lower on all other factors (all Self-Other minus factors and Negativism) than will the vocational failures. In addition to the above, it was hypothesized that since the MSGO is a measure of psychopathology the total patient group will score significantly higher than a matched normal group on all six MSGO factors.

The rationale for the first hypothesis was threefold. First, it was assumed that patients

TABLE 1
SUMMARY OF SIX MSGO FACTORS

MSGO factor	Interpretation
Self-Goal plus	<i>SC-GSC Discrepancy: Culturally typical goals</i> (similar to the commonly used "Self-Ideal" discrepancy, i.e., Factor 1 is related to all self-goal positive discrepancies). High scores on this factor are significantly correlated with neurotic diagnoses.
Self-Goal minus	<i>SC-GSC Discrepancy: Culturally deviant goals</i> (based on rejection of usual or typical goals, i.e., Factor 2 is based on self-goal negative discrepancies). High scores on this factor are significantly correlated with personality disorder and psychotic diagnoses.
Self-Other plus	<i>SC-PRO Discrepancy: Accepting others</i> (others generally overvalue him, do not know him as he "really" is, i.e., Factor 3 is related to self-others' evaluation of self positive discrepancies). High scores on this factor correlate significantly with the neurotic and psychotic depressive diagnoses.
Self-Other minus (Specific)	<i>SC-PRO Discrepancy: Critical others</i> —Specific, interpersonal, affective aspects of self (others generally criticize him as regards his social skills, his emotional stability, his empathy for others, etc., i.e., Factor 4 is based upon self-others' evaluation of self negative discrepancies). High scores on this factor correlate significantly with most psychotic diagnoses.
Self-Other minus (Global)	<i>SC-PRO Discrepancy: Critical others</i> —Global (relates to basic overall evaluations, and he feels that others regard him as "completely" deficient, i.e., Factor 5 is related to self-others' evaluation of self negative discrepancies for the "General" items). High scores on this factor correlate significantly with delusional psychotic diagnoses.
Negativism	<i>Negativistic Response Set: SC-GSC minus on emotional items</i> (Ss who maintain that they wish to be more tense, less self-confident, etc.). High scores on this factor correlate significantly with the psychotic and personality disorder diagnoses.

Note.—SC = self-concept; GSC = goal self-concept; PRO = perceived responses of others.

exhibiting high Self-Goal plus discrepancy, which correlates highly with neurotic psychopathology, would be the group most likely to benefit from psychiatric treatment, and, therefore, would be a group with high likelihood of job obtainment. Second, Miskimins (1967) theorized that persons exhibiting high Self-Goal minus discrepancy would be characterized by the ability to manipulate others, and this manipulative ability can be conceived of as a positive, although not necessarily desirable, factor in obtaining placement. Finally, patients exhibiting severe depression (high Self-Other plus scores) and other psychotic behavior were assumed to be the most difficult to successfully treat, and therefore it was thought that placement would be unlikely.

The basis for the second hypothesis is similar to the first with respect to patients providing high Self-Goal plus discrepancies; that is, of patients obtaining placement, those exhibiting neurotic symptomatology are most likely to achieve vocational success. However, it was thought that high Self-Goal minus patients, although successful in obtaining placement, would fail once their manipulative behaviors became obvious to their employers. The prediction that patients scoring high on the Self-Other plus factor would succeed, once placed, was based on the premise that if depressed patients met with success (placement), this would serve to provide reduction of the discrepancy for this factor (i.e., reduce depression). Related to the placeability predictions, it was assumed that patients who

were negativistic and had considerable difficulty in relating to others (the Self-Other minus factors) would fail if they should happen to obtain placement.

METHOD

Scale

The hypotheses being tested involve the predictive abilities of six "self" factors measured by the MSGO. This scale provides, in addition to information typically gathered in self-concept research, other pertinent self-related data; for this reason, it appears to be the most comprehensive and detailed measurement technique readily available in this area.

Subjects

One hundred and four Ss were divided equally into the four study groups—vocational successes, vocational failures, vocationally unplaceable, and normals. The patient samples were drawn from psychiatric patients referred to the Placement Department at the Fort Logan Mental Health Center, Denver, Colorado. The 26 successes were randomly selected from placement referrals who, at the time of this research, were holding and had held the same job for at least 3 months (averaging over 6 months). The 26 failures were randomly selected from placement referrals who, at the time of this research, were not holding a job and had failed on one or more placements (never remaining on any job over 3 months). The 26 patients classified as unplaceable were randomly selected from persons who, at the time of this research, had no job and had not held a job for over 6 months, during which time Fort Logan's placement services had been available to them. The 26 normals were successfully employed people taken from the community and were matched with the other three groups on age, sex, and level of education. Four age groups were utilized for the matching: 20-30, 31-40, 41-50, and 51 and over. Patient and normal group composition did not differ

TABLE 2
SUMMARY OF MEAN SCORES ON SIX MSGO FACTORS

MSGO factor	Mean scores					
	Normals	Total patients ^a	Placeable ^b	Unplaceable	Successes	Failures
Self-Goal plus	23.731	37.038	39.827	31.462	44.077	35.577
Self-Goal minus	0.423	1.948	2.615	0.615	1.154	4.077
Self-Other plus	7.731	10.846	10.846	10.846	12.577	9.115
Self-Other minus (Specific)	3.692	6.205	5.827	6.962	3.577	8.077
Self-Other minus (Global)	1.385	2.076	1.808	2.615	0.962	2.654
Negativism	0.154	0.858	0.654	1.269	0.500	0.800

^a Consists of successes, failures, and unplaceable patients.

^b Consists of successes and failures.

over 1% for any of the four age groups. With respect to sex, both patient and normal groups were 42% male and 58% female. Three ranges of education were used for the matching: Less than 9 years, 9-12 years, and over 12 years. Patients and normals were within 2% of each other for all three ranges.

Procedure

As part of a routinely administered battery of vocational interest tests, all Ss in the patient group were administered the MSGO soon after admission to Fort Logan (upon referral to Placement). The period of time between MSGO administration and the assessment of vocational adjustment varied from 1 to 12 months. For all patient groups the average elapsed time was approximately 5 months. Normals were tested by a research assistant who went out in the community. The instructions were read to all Ss and example items were carefully completed. The data used in the matching processes described above were obtained from self-report, and in the case of the patient groups, these data were verified through the hospital records.

RESULTS

The mean scores for the four study groups and relevant combinations of these groups on all factors are reported in Table 2. A 1×4 analysis of variance was utilized to test for differences among the means of the four groups on each factor. Five of the six ANOVAs resulted in significant F ratios; Self-Other plus was the only factor which did not differentiate the four groups significantly. Table 3 presents a summary of these results.

In order to test the specific hypotheses it was necessary to run mutually orthogonal comparisons on the presumed relationships, utilizing the within sum of squares from the calculated ANOVAs (see Edwards, 1960).

TABLE 3

SUMMARY OF 1×4 ANALYSES OF VARIANCE

MSGO factor	F	$p <$
Self-Goal plus	4.736	.005
Self-Goal minus	4.600	.005
Self-Other plus	1.674	<i>ns</i>
Self-Other minus (Specific)	4.357	.010
Self-Other minus (Global)	3.300	.025
Negativism	5.068	.001

Note.— $df = 3/100$.

Although the above analysis of the Self-Other plus factor was not significant, it was decided to test the orthogonal comparisons on this factor in order to verify rejection of the hypotheses involving this form of discrepancy. The results of these comparisons supported the hypotheses in all but four instances (see Table 4). Patients and normals were significantly different on all factors except Self-Other minus (Global). Vocationally placeable patients were significantly different from unplaceable patients on the Self-Goal plus and minus factors and on the Negativism factor, while a statistical "tendency" ($p < .100$) supporting the placeability hypothesis was found for the Self-Other minus (Global) discrepancy. The Self-Other plus and minus (Specific) factors did not differentiate the placeable and unplaceable patient groups. For the success versus failure comparisons, significant differences in the predicted direction were found for the Self-Goal minus and both Self-Other minus factors, while the analyses of the

TABLE 4

SUMMARY OF MUTUALLY ORTHOGONAL COMPARISONS

MSGO factor	Normal vs. patient		Unplaceable vs. placeable		Success vs. failure	
	t	p	t	p	t	p
Self-Goal plus	2.958	.001	1.754	.050	1.540	.100
Self-Goal minus	1.666	.050	2.060	.025	2.605	.005
Self-Other plus	1.659	.050	.001	<i>ns</i>	1.506	.100
Self-Other minus (Specific)	1.985	.025	1.014	<i>ns</i>	2.901	.001
Self-Other minus (Global)	1.264	<i>ns</i>	1.389	.100	2.522	.010
Negativism	2.901	.001	2.384	.010	1.037	<i>ns</i>

Self-Goal plus and Self-Other plus discrepancies showed supporting statistical tendencies. Successes and failures were not differentiated by the Negativism factor.

DISCUSSION

The results of this study strongly support the hypothesis that normals would exhibit less internal tension, that is, greater adjustment, than patients on all self variables measured by the MSGO. The only exception to the postulation was the lack of differentiation by Self-Other minus (Global). This lack of difference, however, does not indicate, as other studies have (e.g., Friedman, 1955; Hillson & Worchel, 1957; Ibelle, 1960), that self-concept measures cannot separate normals from delusional patients. Rather, the data in Table 2 suggest that the nondifferentiation was *not* a product of high discrepancy scores for normals, but that the vocational successes low score lowered the total patient mean to a point where no difference between patients and normals was indicated. Indeed, a supplementary orthogonal comparison of normals with the combined vocational failures and unplaceables significantly demonstrates that normals and the two "failure" groups of patients can be differentiated by this aspect of the MSGO ($t = 2.151, p < .025$), while normals and vocational successes cannot be separated on the same factor ($t = 0.630, p = ns$).

With regard to the posthospital vocational aspect of this study, the results support the hypothesized relationships for nine of the 12 comparisons tested. First, the findings of the placeable-unplaceable comparisons will be discussed. The hypothesis that the placeable group would exhibit significantly higher Self-Goal plus discrepancy upon admission than would the unplaceable group was significantly confirmed. As was mentioned earlier, the reason for this confirmation would seem to be that patients exhibiting this form of psychopathology are more likely to be successfully treated than those patients manifesting the self-discrepancies characteristic of more severe forms of psychiatric disability, that is, the psychoses and personality disorders. The significant differentiation between these two groups by the Self-Goal minus factor was not

hypothesized on the basis of ease of symptom remission, rather in this case it was thought that the patient's pathology would contribute to job placement. That is, this discrepancy (self-concept-goal self-concept minus) is highly correlated with the general diagnostic category of personality disorder and more specifically with sociopathic personality disturbance (Miskimins, 1967), a diagnosis characterized, in many cases, by manipulation of others; and it was assumed that this manipulative ability would be an asset during the job-seeking process. The third factor, Self-Other plus, failed to differentiate the placed and not placed groups; in fact, the means on this factor for these two groups are identical. The failure to support the hypothesis that unplaceable patients would manifest significantly higher self-concept-perceived responses of others (SC-PRO) plus discrepancy (depression) than would placeable patients suggests that this discrepancy is not a determining factor in obtaining vocational placement; however, it has been shown to be of some value in predicting vocational success or failure (see Table 4). Self-Other minus (Specific) also failed to separate the placed and not placed. This finding did not support the hypothesis that the unplaceable patient would be characterized by a significantly higher degree of hypersensitivity to criticism of interpersonal and affective aspects of self than would a placeable patient. A further discussion of this finding will be presented with the success-failure comparisons. The fifth MSGO factor, which measures global feelings of criticism from others, did not separate the placed from not placed as sharply as had been expected ($p < .100$), but there was some predictive value indicated. The presence of a statistical tendency for this comparison, rather than significance, will also be illuminated by the discussion of success-failure comparisons. As was expected, the two groups were clearly differentiated by the Negativism factor, that is, the unplaceable patient exhibits significantly more Negativism (a correlate of psychotic behavior) than does the placeable patient. This differentiation suggests that the presence of high discrepancy on this factor is contributory to unplaceability, that is, a person whose emotional makeup

is "negative" will also feel negative toward placement.

The Self-Goal plus factor tends to differentiate vocational successes from failures in the direction hypothesized, that is, there is here indication that the potential success's pathology is characterized by anxiety over attaining acceptable goals, while the failure's illness manifests itself in other forms. The failure group, as hypothesized, scored significantly higher on the Self-Goal minus factor; thereby suggesting that although the presence of this discrepancy is useful for job obtainment, the potential failure's manipulation of others soon becomes obvious and job termination results. The third factor, Self-Other plus, also tends to separate these two groups in the expected direction, thus indicating that this factor is of some predictive value once the patient is placed. As was previously postulated, a possible reason for this success-failure separation would seem to be that if the depressed patient should meet with some success (placement), this would raise his self-concept and thereby reduce the SC-PRO plus discrepancy; and once this partial remission occurs the patient has a higher probability of job success. The two groups' difference on both of the Self-Other minus factors were highly significant, and these differences not only demonstrate the predictive potential of these two factors for job success and failure, but also explain the lack of clear-cut distinction on these factors for the placed-not placed groups. On both of these factors the failures are closely aligned with the not placed in their pathological perception of criticism from others (see Table 2), thereby obscuring any differences between the placed and not placed groups. This lack of distinction between failures and the unplaceable patients is not surprising, as both groups form a larger category of failure. The success-failure comparison for Factor 6, Negativism, does not support the hypothesis, that is, emotional negativism, although predictive of placeability, does not seem to be an indicator of success potential.

The results of this study have clearly demonstrated that the vocational adjustment of psychiatric patients is indicated by their concept of self, as measured by the MSGO. The unplaceable patient is characterized by severe Self-Other minus discrepancies (psychotic disability) and a generally negative emotional set, while the vocationally placeable patient exhibits high Self-Goal plus and minus discrepancies. Within the placeable group, the vocational successes manifest higher Self-Goal plus discrepancy (neurotic psychopathology), while the failures are characterized by high Self-Goal minus (manipulative behavior) and a considerable amount of Self-Other minus tension.

REFERENCES

- EDWARDS, A. L. *Experimental design in psychological research*. New York: Holt, Rinehart & Winston, 1960.
- FRIEDMAN, I. Phenomenal, ideal and projected conceptions of self. *Journal of Abnormal and Social Psychology*, 1955, **51**, 611-615.
- HILLSON, J. S., & WORCHEL, P. Self concept and defensive behavior in the maladjusted. *Journal of Consulting Psychology*, 1957, **21**, 83-88.
- IBELLE, B. P. *Discrepancies between self-concepts and ideal self-concepts in paranoid schizophrenics and normals*. (Doctoral dissertation, University of Connecticut) Ann Arbor, Mich.: University Microfilms, 1960. No. 60-5237.
- MISKIMINS, R. W. *The concept of self and psychopathology*. (Doctoral dissertation, University of Colorado) Ann Arbor, Mich.: University Microfilms, 1967. No. 67-422.
- MISKIMINS, R. W. *Manual: MSGO discrepancy scale*. Ft. Collins, Colo.: Rocky Mountain Behavioral Science Institute, 1968.
- MISKIMINS, R. W., & SIMMONS, W. L. Goal preference as a variable in involutional psychosis. *Journal of Consulting Psychology*, 1966, **30**, 73-77.
- ROGERS, C. R., & DYMOND, R. F. (Eds.) *Psychotherapy and personality change: Co-ordinated studies in the client-centered approach*. Chicago: University of Chicago Press, 1954.
- SARBIN, T. R., & ROSENBERG, B. G. Contributions to role-taking theory: IV. A method for obtaining a qualitative estimate of the self. *Journal of Social Psychology*, 1955, **42**, 71-81.

(Received March 22, 1968)

PERSONALITY AND PERFORMANCE CHARACTERISTICS ASSOCIATED WITH CIGARETTE SMOKING AMONG COLLEGE FRESHMEN¹

DONALD J. VELDMAN² AND OLIVER H. BOWN

The University of Texas at Austin

Habitual cigarette smokers were located in a group of 2321 university freshmen. Smoking was related to both sex and social class. Matched samples of 401 smokers and nonsmokers were compared on a series of 70 dependent variables. Smokers were significantly lower in cognitive ability, high school and college grades, two self-report health indices, self-esteem, work-motivation, and attitudes toward men. The smokers scored higher on semiprojective measures of orality, social confidence, and acceptance of ambiguity. The data were most consistent with the stimulus hunger theory of smoking motivation.

In 1964 the Office of the United States Surgeon General issued an extensive report, *Smoking and Health*, which included correlative evidence linking cigarette smoking with a number of physical diseases. The report also considered the question of personality traits associated with the smoking habit. The authors summarized the available literature as follows:

Despite the individual deficiencies of many of the studies, despite the great diversity in conceptualization and research methods used, and despite certain discrepancies in reported findings, the presence of some comparability between them and the relative consistency of findings lend support to the existence of a relationship between the smoking habit and a personality configuration that is vaguely described as "neurotic" [United States Public Health Service, p. 367].

In a review of the literature on personality and smoking, Matarazzo and Saslow (1960) mentioned large-scale statistical summaries in the 1950s which indicated that about 46% of the males and 31% of the females in the age range 18-24 smoked cigarettes. Their own data suggested that smoking was related to income, anxiety, psychological tension, and

self-reports of psychological symptoms, but not to social class level. Among the earlier studies they reviewed was one by McArthur, Waldron, and Dickinson (1958) where smoking was more common among lower-class public school students.

In 1961 Salber and MacMahon studied all high school students in Newton, Massachusetts, and found that the percentages of smokers were greater at the lower social class levels. Horn (Horn, Courts, Taylor, & Solomon (1959) also found that smoking was more common among students from lower social class levels (as measured by fathers' education); the association was strongest in Grade 9, but decreased steadily to become negligible in Grade 12. The authors suggested that this phenomenon was due to lower-class students beginning to smoke earlier, but the same effect would be produced if lower-class smokers were more likely to drop out of school before graduation.

Another major topic of many studies of personality and smoking has been a hypothetical syndrome of personality traits suggesting character disorder rather than neurosis as the "pathology" underlying the smoking habit. Straits and Sechrest (1963) found that male smokers scored significantly higher on the Psychopathic Deviate (*Pd*) scale and lower on the Social Responsibility (*Sr*) scale of the MMPI; other data indicated that smokers were more "chance oriented." Schubert (1959) also found a *Pd* scale difference with college freshmen. Stewart and Livson (1966) noted that smokers scored

¹ This research was supported in part by the Excellence Fund of the University of Texas, Grant No. 06823 from the National Institute of Mental Health (Computer Analysis of Personality), and by the Research and Development Center for Teacher Education, United States Office of Education Contract OE6-10-108.

² Requests for reprints should be mailed to Donald J. Veldman, R&D Center-Education, University of Texas, Drawer J, University Station, Austin, Texas 78712.

lower than nonsmokers on the Socialization scale of the CPI. Their other data strongly suggested that the "rebelliousness" of the smokers antedated the habit by many years. Carney (1967) obtained similar results with the same measure, and also found a positive correlation between smoking and the Achievement Orientation (Ao) scale of the CPI. Jacobs (Jacobs, Knapp, Anderson, Karush, Meissner, & Richman, 1965) described the personality trait syndrome associated with smoking in terms of "oral frustration, oral impulsivity, defiance of authority, and living more in the present than in the future."

The evidence from self-report personality measures is not entirely uniform. Weatherly (1965) found no differences at all between smokers and nonsmokers on the scales of the Edwards Personal Preference Schedule. Since this instrument was designed to control the influence of social desirability response set, he suggested that the effects reported in other studies may only indicate that smokers are simply more willing to acknowledge socially undesirable characteristics. When Jacobs and his colleagues (Jacobs, Knapp, Anderson, Karush, Meissner, & Richman, 1966) attempted a direct test of the response-set hypothesis, they found clear attitudinal differences between smokers and nonsmokers, but no evidence of a generalized response set that could account for them. Particularly interesting is their summary of the smoker syndrome as reflecting "... a neurotic image depreciatory to both self and family."

The characterization of the smoker as "neurotic" appears now and then in the literature. Lilienfeld (1959), for instance, found smokers to be more neurotic, and to have been hospitalized more frequently than nonsmokers. Eysenck (1960, 1963), beginning with the notion that *both* smoking and carcinoma are resultants of a third genotypic factor, hypothesized that smokers would be more extraverted, less rigid, and more neurotic. Data from very large samples of males clearly confirmed the smokers' extraversion, mildly supported their flexibility, but failed entirely to justify the term "neurotic." Eysenck has suggested that both smoking and overeating are symptoms of "stimulus hunger," and that when smoking is stopped,

it is replaced by overeating—a well-known phenomenon. Carney (1967), in the study mentioned earlier, found relationships between sex chromatin, masculinity of body type, and smoking among college students, and suggested the same "third factor" link between smoking and cancer as a possibility. A recent study by Pumroy (1967) of male college freshmen enrolled in physical education indicated that amount of cigarette smoking was negatively related to grade-point average and positively to the probability of dropping out of school. These relationships were interpreted as due to the smokers' supposed antagonism toward parents and others in authority.

Finally, two studies by Baer (1966a, 1966b) with college males produced rather surprising results that deserve further investigation. He found that cigarette smokers (former and current) obtained significantly lower scores on the mathematical portion of the College Entrance Examination Board Scholastic Aptitude Test. He also failed to discriminate smokers from nonsmokers on the basis of height, weight, or a ponderal index derived from these two variables.

The analyses to be reported in the remainder of this paper were designed to resolve certain ambiguities in many of the earlier studies, and to confirm the results of previous investigations where possible within the limits of our data. In particular, the variables of sex and social class were considered important enough to enter into the selection of the Ss, and were treated subsequently as independent variables in the design so that their interactions with smoking behavior could be assessed.

INDEPENDENT VARIABLES

The basic data from which the present samples were drawn were collected from a total of 2321 freshmen at the University of Texas as part of an institutional research project. This sample constituted the majority of the incoming freshman class during the fall, 1963, semester.

Three paper/pencil instruments were administered: (a) the Bown *Self-Report Inventory*, (b) an Autobiographical Information (AI) form, and (c) the Veldman-Peck *One-Word Sentence Completion* (Form 63). The

definition of a "smoker" was based on a single item from the autobiographical form; at the end of a health status checklist was the following:

— smoke cigarettes every day

Anyone who checked this item was classified as a smoker. Although considerable variation in the intensity of the habit is possible here, habituation can be safely assumed from a positive response.

Another item on the autobiographical form asked *S* to indicate the extent of his father's education. This variable was used as an operational definition of socioeconomic class. Three levels were defined: (a) high school only, (b) college only, and (c) graduate degree.

Of the original 2321 *Ss*, 100 failed to indicate their fathers' educational levels. Table 1 shows the distribution of the remaining 2221 *Ss* by smoking, sex, and social class level. By use of the three sets of marginal frequencies, expected values for each of these 12 cells were computed, and the total chi-square was found to be highly significant ($\chi^2 = 34.31$, $df = 11$, $p = .0005$). Two-way chi-squares were then computed for all combinations of the three variables.

The chi-square value for the association between sex and smoking was statistically significant ($\chi^2 = 4.58$, $df = 1$, $p = .03$); males (19.5%) were more likely to smoke than were females (16%). The chi-square value for the association between social class and smoking was also significant ($\chi^2 = 13.98$,

TABLE 2

DISTRIBUTION OF SELECTED SAMPLES BY SEX
AND FATHERS' EDUCATION

Fathers' education	Subject sex		
	Males	Females	Totals
High school	90	41	131
College	83	58	141
Postgraduate	81	48	129
Totals	254	147	401

$df = 2$, $p = .001$); smoking was much more common at the highest social class level (23.6%), and roughly equal (16.2%) at the lower levels. The chi-square value for the association of sex and social class level was also significant ($\chi^2 = 14.99$, $df = 2$, $p = .0009$); the proportion of males at the lowest social class level (63.7%) exceeded those at the higher levels (55.4% and 55.9%). The residual chi-square value (Sex \times Smoking \times Social Class) did not approach statistical significance.

In order to avoid biases in cell variance due to very disproportionate subsample sizes, further *S* selection was carried out to equate the groups of smokers and nonsmokers. For each of the 401 smokers, a nonsmoker of the same sex and social class level was randomly selected. The resulting distributions of *Ss* by sex and social class are shown in Table 2. All statistical analyses reported subsequently were based on this selected group of 802 students.

All but two of the 70 dependent variables studied were continuously distributed, and were subjected to a three-factor analysis of variance design (Veldman, 1967), in which the three main effects (smoking, sex, and social class) and four interactions were all tested against an averaged within-cells error term. Two categorical variables, birth-order position and future vocational preference, were analyzed by means of the chi-square technique.

DEPENDENT VARIABLES

Cognitive, Somatic, and Social Variables

Data for these variables were obtained from the AI form and from central records of the University.

TABLE 1

DISTRIBUTION OF STUDENTS BY SMOKING, SEX,
AND SOCIAL CLASS LEVEL

Smokers	Social class			Totals
	HS	C	G	
Males				
No	425	397	225	1047
Yes	90	83	81	254
Totals	515	480	306	1301
Females				
No	252	328	193	773
Yes	41	58	48	147
Totals	293	386	241	920

*Social class estimated from father's education: HS = high school, C = college, G = postgraduate degree.

Verbal ability and quantitative ability. These scores were from the Scholastic Aptitude Test.

High school quarter. This information was obtained from the files of the Testing and Counseling Center. Although a low number signifies a high rank, subsequent discussions will reflect the level of performance implied.

Freshman grade-point average (GPA). This score was determined from university records, and covers only the initial semester during which enrollment and testing took place.

Confidence in college. This score was based on a three-choice item from the AI form which asked for an estimate of certainty about eventually graduating from college. Again, a low score signifies certainty, but the discussion will reflect degree of confidence.

Weight and height. These data were obtained from the AI form.

Health index. A six-item checklist was provided on the AI to register various types of physical disorders. Since the sum of items checked was used as the score, a low value signifies health, and subsequent discussions will reflect this interpretation.

High school social activity. The AI included space for S to list extracurricular activities. These were simply counted; no attempt was made to weight various types of activity.

Number of siblings. This score was based on information from the AI form, as was the next variable.

Birth order position. Three categories were employed: (a) only child, (b) firstborn, and (c) later born.

Future occupational choice. Data were obtained from the Testing and Counseling Center, and the most frequently named categories were used (see Tables 4 and 5).

Self-Report Inventory Variables

The *Self-Report Inventory* (Bown, 1961; Bown & Richek, 1967) consists of 48 statements, which are scored for eight subscales and a total of all items. The item statements are relatively straightforward self-reports of the respondent's feelings toward various aspects of his phenomenal world. Internal consistency coefficients (α) are included with the scale descriptions that follow.

Self. ($\alpha = .78$) Acceptance, liking, valuing of oneself.

Others. ($\alpha = .65$) Attitude toward peers; the importance of satisfactory interpersonal relationships to own sense of well-being.

Children. ($\alpha = .85$) Attitude toward children; satisfaction derived from relationships with them.

Authority. ($\alpha = .53$) Attitude toward non-familial authority figures.

Work. ($\alpha = .70$) Valuing of work and accomplishment in terms of intrinsic satisfaction or self-enhancement.

Reality. ($\alpha = .28$) Acceptance of life as a process (including death); positive existential adjustment.

Parents. ($\alpha = .84$) Attitude toward parents; importance of relationships with them to own sense of well-being.

Hope. ($\alpha = .66$) Optimistic anticipation of the future.

Total. ($\alpha = .87$) Simple sum of the eight scales: the general positiveness of the respondent's view of his phenomenological world.

Detailed explication of the scoring procedures and norms are available from the authors (Bown & Veldman, 1967). In order to obtain more sensitive measures of attitudes toward Others, Children, Authority, and Parents, the Self score was subtracted from each of these scale scores to yield four additional measures. Both the Self scores and the scores for attitudes toward others reflect two components, the "true" attitude and a general response set. By subtracting the Self score from those for other significant figures, the common response set is canceled and the "true" attitude is also made relative to S's own level of self-esteem.

Sentence Completion Variables

As part of a research program concerning computer-based psychological assessment,^a a 36-item sentence completion form was included in the test battery. The Ss were restricted to single-word responses, which were transferred to punch cards. An experimental scoring procedure employing preclassified word

^aGrant 06823 from the National Institute of Mental Health, titled "Computer Analysis of Personality," codirectors Robert F. Peck and Donald J. Veldman.

roots was then used to obtain quantitative indices on 40 variables. Most of the basic procedures are described elsewhere (Veldman, 1967).

1. Response length. Average number of characters per response.
2. Response variation. Number of generic response types used.
3. Consistency index. Number of identical responses to four paired stems.
4. Pronoun responses. (e.g., "*He* makes me angry.")
5. Evasive responses. Stem-repetitive words and other noninformatives.
6. Proper names. (e.g., "*John* makes me feel sad.")
7. Nonresponses. Unanswered items.
8. Self references. (e.g., "The average person is *me*.")
9. Age responses. (e.g., "I am 18.")
10. Orality. References to food, eating, drinking, and smoking.
11. Money. References to acquisition, use, or lack of money.
12. Ideology. References to politics or religion.
13. Commonality. Ordinarity (high) versus uniqueness (low) in response to a variety of self-descriptive stems.
14. Somatic self-esteem. Health, strength, and attractiveness.
15. Social self-esteem. Extraversion versus shy or withdrawn.
16. Cognitive self-esteem. Intellectual depth and acuity.
17. Performance self-esteem. Success or improvement.
18. General self-esteem. Mostly the sum of scores 14-17.
19. Emotional stability. ("My emotions are _____.") Controlled, mild versus erratic, intense.
20. Character. Morality, stability, responsibility.
21. Impulse acceptance. ("When an animal is wild, it is _____.") Approval, acceptance versus fear, rejection.
22. General mood. Happiness and satisfaction.
23. Optimism. Expectations for the future.

24. Certainty. References to decisiveness versus doubt.

25. Self-confidence. Level of conscious anxiety.

26. Ambiguity acceptance. ("Darkness is _____.") Acceptance, pleasure versus rejection, fear, depression.

27. Stress resistance. ("I _____ when put under pressure.") Active coping versus anxiety, quitting, or rebellion.

28. Academic attitude. Cathexis versus rejection of school and studies.

29. General motivation. Ambition, interest, and effort.

30. Attitude toward mother. ("My mother is _____.")

31. Attitude toward father. ("My father is _____.")

32. Attitude toward family. Mostly the sum of scores 30 and 31.

33. Attitude toward men. ("Most men are _____.")

34. Men toward women. ("Men often _____ women.")

35. Attitude toward women. ("Most women are _____.")

36. Women toward men. ("Women often _____ men.")

37. Heterosexuality. Cathexis of opposite sex and marriage. Derived largely from scores 33-36, considering sex of S.

38. Attitude toward average person. ("The average person is _____.")

39. General interpersonal. Friendliness, kindness, courtesy.

40. Unique responses. Count of responses that could not be located by the word-root system. (Less than 2% of the total.)

As noted in the definitions above, a number of the scores are based on responses to single stems, while others are derived from particular aspects of the entire protocol. No reliability data are available at present on this system, but Peck, Menaker, and Veldman (1966) reported very satisfactory agreement between ratings by human judges and scores from an earlier computer-based system.

RESULTS

For ease of communication, the results of the statistical analyses will be presented in

TABLE 3

RESULTS OF ANALYSES OF COGNITIVE, SOMATIC, AND SOCIAL VARIABLES

Dependent variable	Smoking		Sex		Social class	
	<i>p</i> level	Direction	<i>p</i> level	Direction	<i>p</i> level	Direction
Verbal ability	.02	NS>S	—	—	.03	C>G>HS
Quantitative ability	.0001	NS>S	<.0001	M>F	.004	C>G>HS
High school quarter	<.0001	NS>S	.003	F>M	—	—
Freshman GPA	<.0001	NS>S	—	—	.003	G>C>HS
Confidence in college	.01	NS>S	<.0001	M>F	.0006	G>C>HS
Weight	—	—	<.0001	M>F	—	—
Height	—	—	<.0001	M>F	—	—
Health index	.001	NS>S	<.0001	M>F	—	—
High school social activity ^a	.003	NS>S	<.0001	F>M	—	—
Number of siblings ^a	—	—	.02	M>F	—	—

Note.—Social class estimated from father's education: HS = high school, C = college, G = postgraduate degree.

^a Also yielded a significant Sex X Smoking interaction.

three groups: (a) the cognitive, social, and somatic variables, (b) the scale scores from the *Self-Report Inventory*, and (c) the *One-Word Sentence Completion* variables. Rather than attempt to report these analyses in full, we have summarized the significant ($p < .05$) results in Tables 3, 7, and 8. Although seven *F* ratios were computed in each analysis, only nine significant interactions occurred. Since they all involved the smoking variable, they will be described below.

Cognitive, Somatic, and Social Variables

Nonsmokers scored higher on both the verbal and quantitative sections of the college entrance exams, but the differences were much greater for the quantitative variable (543 vs. 515) than for the verbal (527 vs. 512). Nonsmokers ranked higher in their high school

classes (1.32 vs. 1.74), and also obtained higher grades during their first college semester (1.32 vs. .90). As might be expected from the performance differences, nonsmokers indicated more confidence (1.83) that they would eventually graduate than did smokers (1.97).

Smoking was not significantly related either to height or to weight, but smokers reported more physical disorders (.55) than did the nonsmokers (.36). Among the males, smokers reported fewer (2.9) social activities in high school than did nonsmokers (3.8), but the difference between the female groups was negligible. With regard to family size, nonsmokers of both sexes were equivalent; male smokers, however, came from somewhat larger families (2.0) than did female smokers (1.6).

Table 4 contains the frequency distribution

TABLE 4

DISTRIBUTION OF SAMPLE BY SMOKING, SEX, SOCIAL CLASS, AND BIRTH-ORDER POSITION

Birth-order position	High school social class		College social class		Postgraduate social class	
	Nonsmokers	Smokers	Nonsmokers	Smokers	Nonsmokers	Smokers
Males						
Only child	21	8	7	9	10	6
Firstborn	30	31	45	31	45	40
Later born	39	51	31	43	26	35
Females						
Only child	7	9	7	3	5	9
Firstborn	15	16	25	19	17	22
Later born	19	16	26	36	26	17

of the sample by sex, social class, smoking, and birth-order position. A three-way chi-square analysis was computed separately for males and females. Birth-order was related to smoking among males ($\chi^2 = 9.99$, $df = 2$, $p = .007$); the proportion of smokers was much smaller among firstborn and only children than among the later born. The three-way interaction of smoking, social class, and birth order was not significant. Among females virtually no differences in proportions of smokers appeared among the birth-order groups.

Tables 5 and 6 contain the frequency distributions among the most commonly designated occupational preferences in the male and female samples. Although the chi-square values for these tables did not approach statistical significance, some rather striking consistencies appear. In both sex samples the Fine Arts group contains twice as many smokers as nonsmokers. In the male sample the Engineering group contains a predominance of nonsmokers, which is consistent with the higher quantitative ability of the nonsmoking Ss.

Self-Report Inventory Scales

The results summarized in Table 7 clearly evidence substantial differences in attitudes between the groups of smokers and nonsmokers. Smokers indicated a lower level of self-esteem than nonsmokers. Although the absolute levels of attitudes toward others did not differ, when these scores were made relative to self-esteem, the smokers appeared to be more positive than nonsmokers. Attitudes toward children were equivalent in the two

TABLE 6

DISTRIBUTION OF FEMALE SAMPLE BY SMOKING AND OCCUPATIONAL PREFERENCE

Occupation	Nonsmokers	Smokers	Percentage
Fine arts	11	22	67%
Business	10	10	50%
Medical-dental	20	21	51%
Social sciences	22	25	53%
Elementary education	17	15	47%
Secondary education	28	20	42%
Other and undecided	39	34	47%
Totals	147	147	

groups. In absolute terms, male smokers indicated more negative attitudes toward non-parental authority figures, but this effect disappeared when the scores were adjusted for level of self-esteem. Smokers indicated more negative attitudes toward work and reality. Smokers reported more negative attitudes toward parents, but the three-way interaction indicated that this was due largely to the extremely low scores of the lower social class, female smokers. When adjusted for level of self-esteem, no significant effects were found. No significant differences in optimism about the future appeared. Smokers obtained a lower average score on the total of all items, but a significant Sex \times Smoking interaction revealed much larger differences between smokers and nonsmokers among males ($M = 131$ vs. 123) than among females ($M = 135$ vs. 133). The male smokers exhibited a strong general tendency toward negative characterization of their own attitudes.

Sentence Completion Variables

A number of intriguing results appear in Table 8, but due to the experimental status of the scoring system that was used to quantify S's verbal behavior, the interpretations placed upon the observed effects must be considered as tentative, and in need of confirmation from other sources. In some cases we will be able to offer such confirmatory evidence from other data analyzed in the present investigation, but other findings must await support from further research.

TABLE 5

DISTRIBUTION OF MALE SAMPLE BY SMOKING AND OCCUPATIONAL PREFERENCE

Occupation	Nonsmokers	Smokers	Percentage
Fine arts	9	17	65%
Business	32	32	50%
Medical-dental	49	42	46%
Engineering	64	49	43%
Law, government	38	39	51%
Research, college teaching	18	26	59%
Other and undecided	44	49	53%
Totals	254	254	

TABLE 7

RESULTS OF ANALYSES OF RAW AND ADJUSTED SELF-REPORT INVENTORY SCALE SCORES

SRI variable (Raw scores)	Smoking		Sex		Social class	
	p level	Direction	p level	Direction	p level	Direction
Self	.01	NS>S	—	—	—	—
Others	—	—	.0004	F>M	—	—
Children	—	—	<.0001	F>M	—	—
Authority ^a	.03	NS>S	.002	F>M	—	—
Work	<.0001	NS>S	.01	F>M	—	—
Reality	.03	NS>S	.004	F>M	—	—
Parents ^b	.0004	NS>S	—	—	—	—
Hope	—	—	—	—	—	—
Total ^a	.0009	NS>S	<.0001	F>M	—	—
Others—Self	.002	S>NS	—	—	—	—
Children—Self	—	—	<.0001	F>M	.03	HS>G>C
Authority—Self	—	—	—	—	—	—
Parents—Self	—	—	—	—	—	—

^a Sex × Smoking interaction.^b Triple interaction.

Nonsmokers used significantly longer response words than did smokers. This finding is consistent with the greater verbal ability of nonsmokers reported earlier.

The significant interaction of Smoking × Social Class with regard to the use of proper names is difficult to explain. Among the lowest and highest social class levels, smokers used more proper names, while the reverse was true for the middle social class level. The triple interaction for nonresponses (blank) was even less interpretable.

Age responses were more common among nonsmokers. Although this could reflect concern with a "maturity image," such an interpretation is speculative at best.

The highly significant tendency for smokers to give more "oral" responses cannot be explained by the fact that the term "smoking" was included in this category; the frequency of this response was far too small. Unfortunately, we can offer no corroborative evidence from the present data for this classic, face-valid theoretical motivation for smoking.

The significantly more frequent ideological references by nonsmokers, and their higher scores on the "character" variable, suggest that moral considerations may play a stronger part in the inhibition of smoking than has been indicated in previous studies.

The higher scores of nonsmokers with regard to somatic self-esteem are consistent with their higher scores on the health index described earlier.

Although the main effect for social self-esteem appears to favor smokers in general, inspection of the means for the Sex × Smoking interaction indicates that this is due almost entirely to extremely low scores for female nonsmokers on this variable. Essentially the same pattern characterizes the analysis of the general self-esteem scores.

The significantly higher scores on the Optimism scale for nonsmokers did not appear with the Hope scale of the Self-Report Inventory, although both measures supposedly concern expectations for the future.

A very interesting difference between smokers and nonsmokers appears on the Ambiguity Acceptance scale, which is derived entirely from the item, "Darkness is _____." Smokers apparently feel less threatened, or more positive toward this stimulus.

Nonsmokers of both sexes indicate more positive attitudes toward men, more positive expectations for the treatment of women by men, and a more positive attitude toward marriage in particular. The complementary scales concerning attitudes toward women, however, do not differentiate the smokers and

TABLE 8
RESULTS OF ANALYSES OF SENTENCE COMPLETION VARIABLES

One-word sentence completion variable	Smoking		Sex		Social class	
	<i>p</i>	Direction	<i>p</i>	Direction	<i>p</i>	Direction
1. Response length	.03	NS>S	—	—	—	—
2. Response variation	—	—	.006	F>M	—	—
3. Consistency index	—	—	.04	M>F	—	—
4. Pronoun responses	—	—	—	—	—	—
5. Evasive responses	—	—	.004	M>F	—	—
6. Proper names ^b	—	—	—	—	—	—
7. Nonresponses ^c	—	—	—	—	—	—
8. Self references	—	—	—	—	—	—
9. Age responses	.02	NS>S	—	—	—	—
10. Orality	.0001	S>NS	.005	M>F	—	—
11. Money	—	—	<.0001	M>F	—	—
12. Ideology	.05	NS>S	—	—	—	—
13. Commonality	—	—	.02	M>F	—	—
14. Somatic self-esteem	.04	NS>S	—	—	—	—
15. Social self-esteem ^a	.002	S>NS	.04	M>F	.05	G>HS>C
16. Cognitive self-esteem	—	—	.003	M>F	—	—
17. Performance self-esteem	—	—	.05	F>M	—	—
18. General self-esteem ^a	—	—	.01	M>F	—	—
19. Emotional stability	—	—	.006	M>F	.04	G>C>HS
20. Character	.0008	NS>S	.005	F>M	—	—
21. Impulse acceptance	—	—	—	—	—	—
22. General mood	—	—	—	—	—	—
23. Optimism	.005	NS>S	—	—	—	—
24. Certainty	—	—	—	—	.02	G>HS>C
25. Self-confidence	—	—	.006	M>F	.02	C=G>HS
26. Ambiguity acceptance	.03	S>NS	—	—	.002	G=C>HS
27. Stress resistance	—	—	.008	M>F	—	—
28. Academic attitude	—	—	—	—	—	—
29. General motivation	—	—	—	—	—	—
30. Attitude to mother	—	—	—	—	—	—
31. Attitude to father	—	—	—	—	—	—
32. Attitude to family	—	—	—	—	—	—
33. Attitude to men	.02	NS>S	.003	F>M	.001	C=HS>G
34. Men toward women	.04	NS>S	.02	M>F	—	—
35. Attitude to women	—	—	.0001	M>F	—	—
36. Women toward men	—	—	—	—	—	—
37. Heterosexuality	.01	NS>S	.0003	F>M	—	—
38. Attitude to average person	—	—	—	—	—	—
39. General interpersonal	—	—	<.0001	F>M	—	—
40. Unique responses	—	—	—	—	—	—

^a Smoking X Sex interaction.

^b Smoking X Social Class interaction.

^c Triple interaction.

nonsmokers. The implications of these results are not at all apparent.

CONCLUSIONS

Some of the conclusions stated below are clearly supported by the present data, while others are offered as speculations deserving further research. In any event, these generali-

zations must be limited by consideration of the population sampled—freshmen at a large university. It should also be noted that any inferences to be drawn from these findings regarding the etiology of the smoking habit must be tempered by the fact that a major influence—the smoking habits of the parents—could not be included in the analyses reported here. With these reservations, the results of

the present investigation may be summarized as follows.

1. Smoking was more common among males than among females. Although relatively few of the differences between smokers and non-smokers needed qualification due to a Sex \times Smoking interaction, the differences due to smoking were generally greater among males than among females.

2. Smoking was more common among students whose fathers held graduate degrees than at the lower social class levels. Generally speaking, very few interaction effects between smoking and social class were observed, however.

3. Although smokers did not tend to come from larger families, among the males only the smokers were much more likely to be later-born than firstborn or only children.

4. Smokers reported more physical disorders and described themselves more negatively with regard to somatic characteristics. They did not differ from nonsmokers in either height or weight, however.

5. Smokers obtained lower academic aptitude scores than did nonsmokers, particularly in the quantitative area. This may explain the fact that they also obtained lower grades in high school and in their first semester at the university.

6. Although male smokers reported involvement in fewer high school social activities, female nonsmokers indicated less social involvement and self-assurance at the time of testing. Various measures of attitudes toward peers yielded equivocal results.

7. The negative attitudes toward parents and other authority figures that are often assumed to characterize the smoker did not emerge clearly in our data. One analysis did suggest that the lower social class girls who smoked held very negative attitudes toward their parents.

8. Nonsmokers revealed a strong general tendency to respond more positively to the self-report attitude measures, and also indicated a certain pride in their own moral character.

9. A few results taken together seem to support the "stimulus hunger" theory of smoking

motivation. The "orality" of the smoking habit is supported by more frequent responses by smokers with content concerning eating and drinking. They also reacted more positively to the stimulus "darkness," when presented as an incomplete sentence.

10. In general, our results do not appear to justify the term "neurotic" as a characterization of the smokers' personality syndrome—if indeed such a pattern exists anywhere but in the minds of those who seek to discourage the habit by means of propaganda.

REFERENCES

- BAER, D. J. Height, weight, and ponderal index of college male smokers and non-smokers. *Journal of Psychology*, 1966, **64**, 101-105. (a)
- BAER, D. J. Scholastic aptitude and smoking attitude and behavior of college males. *Journal of Psychology*, 1966, **64**, 63-68. (b)
- BOWN, O. H. *Self-Report Inventory*. University of Texas at Austin, 1961.
- BOWN, O. H., & RICHEL, H. G. The Bown Self-Report Inventory (SRI): A quick screening instrument for mental health professionals. *Comprehensive Psychiatry*, 1967, **8**, 45-52.
- BOWN, O. H., & VELDMAN, D. J. Scoring procedure and college freshman norms for the *Self-Report Inventory*. Research Methodology Monograph No. 1, R & D Center for Teacher Education, University of Texas at Austin, 1967.
- CARNEY, R. E. Sex chromatin, body masculinity, achievement motivation and smoking behavior. *Psychological Reports*, 1967, **20**, 859-866.
- EYSENCK, H. J., TARRANT, M., & WOOLF, M. Smoking and personality. *British Medical Journal*, 1960, **1**, 1456-1460.
- EYSENCK, H. J. Personality and cigarette smoking. *Life Sciences*, 1963, **3**, 777-792.
- HORN, D., COURTS, F. A., TAYLOR, R. M., & SOLOMON, E. S. Cigarette smoking among high school students. *American Journal of Public Health*, 1959, **49**, 1497-1511.
- JACOBS, M. A., KNAPP, P. H., ANDERSON, L. S., KARUSH, N., MEISSNER, R., & RICHMAN, S. J. Relationship of oral frustration factors with heavy cigarette smoking in males. *Journal of Nervous and Mental Disease*, 1965, **141**, 161-171.
- JACOBS, M. A., KNAPP, P. H., ANDERSON, L. S., KARUSH, N., MEISSNER, R., & RICHMAN, S. J. Orality, impulsivity and cigarette smoking in men: Further findings in support of a theory. *Journal of Nervous and Mental Disease*, 1966, **143**, 207-219.
- LILIENTHAL, A. M. Emotional and other selected characteristics of cigarette smokers and non-smokers as related to epidemiological studies of lung cancer and other diseases. *Journal of the National Cancer Institute*, 1959, **22**, 259-282.

- MATARAZZO, J. D., & SASLOW, G. Psychological and related characteristics of smokers and non-smokers. *Psychological Bulletin*, 1960, 57, 493-513.
- MCARTHUR, C., WALDRON, E., & DICKINSON, J. The psychology of smoking. *Journal of Abnormal and Social Psychology*, 1958, 56, 267-275.
- PECK, R. F., MENAKER, S., & VELDMAN, D. J. Computer analysis of sentence completion data. *American Psychologist*, 1966, 21, 678. (Abstract)
- PUMROY, D. K. Cigarette smoking and academic achievement. *Journal of General Psychology*, 1967, 77, 31-34.
- SALBER, E. J., & MACMAHON, B. Cigarette smoking among high school students related to social class and parental smoking habits. *American Journal of Public Health*, 1961, 51, 1780-1789.
- SCHUBERT, D. S. P. Personality correlates of cigarette smoking among college students. *Journal of Consulting Psychology*, 1959, 23, 376.
- STEWART, L., & LIVSON, N. Smoking and rebelliousness: A longitudinal study from childhood to maturity. *Journal of Consulting Psychology*, 1966, 30, 225-229.
- STRAITS, B. C., & SECHREST, L. Further support of some findings about the characteristics of smokers and non-smokers. *Journal of Consulting Psychology*, 1963, 27, 282.
- UNITED STATES PUBLIC HEALTH SERVICE. *Smoking and health*. Public Health Service Publication No. 1103. Washington, D. C.: USPHS, 1964.
- VELDMAN, D. J. A study of freshmen: VQ analyses. Personality Research Center, University of Texas at Austin, 1966. (Mimeo.)
- VELDMAN, D. J. *Fortran programming for the behavioral sciences*. New York: Holt, Rinehart & Winston, 1967.
- WEATHERLY, D. Some personality correlates of the ability to stop smoking cigarettes. *Journal of Consulting Psychology*, 1965, 29, 483-485.

(Received March 22, 1968)

NOTES AND COMMENTS

INTERRELATIONSHIPS OF SEVERAL PROCESS-REACTIVE MEASURES¹

CHARLES G. WATSON

AND

PATRICK E. LOGUE

Veterans Administration Hospital,
St. Cloud, Minnesota

Veterans Administration Hospital,
Coral Gables, Florida

8 previously used and potential process-reactive measures were intercorrelated and factored by the principal-factor solution. 2 factors were rotated. The 1st factor, consisting of marital status and the Phillips, Elgin, and Ullmann-Giovannoni scales, consisted of traditional process-reactive measures. The variables making up the 2nd factor do not appear to be appropriate process-reactive indices. Implications of the findings are briefly discussed.

Research on the process-reactive dimension has proliferated in recent years as psychologists have used the concept in an effort to bring order to the schizophrenia literature. A questionable by-product of this activity has been the disorderly development of a rash of process-reactive scales and definitions. The operational definitions which have appeared in the literature include the Elgin Prognostic Scale (Becker, 1955, 1956, 1959; Wittman, 1941), the Phillips Premorbid Scale (Phillips, 1953), the Kantor Scale (Kantor, Wallner, & Winder, 1953), the MMPI Ego Strength (*Es*) scale (Herron, 1962), the PR questionnaire (Ullmann & Giovannoni, 1964), Rorschach scores (Kantor & Herron, 1966), an unsuccessful MMPI scale (Johnson & Holmes, 1967), and marital status. Moreover, other measures, such as length of previous psychiatric hospitalization, might also be proposed as appropriate process-reactive indices.

To the extent that these measures may be independent, psychologists run the risk of developing separate—and possibly conflicting—bodies of process-reactive literature. The purpose of the present project was to study the interrelationships of eight previously used and potential process-reactive measures by factor analysis to determine whether they reflect the same dimension.

METHOD

Ninety-four schizophrenic patients at the St. Cloud Veterans Administration Hospital served as Ss. All

¹ Carried out within the framework of activities supported by the Veterans Administration Program Evaluation Staff. The authors are indebted to Raymond Collier, Theodore W. Lorei, and Jack Collier for help with the statistical analyses and to Vi Kohorst, Catherine Tidd, and Kathleen Schelonka for clerical support.

were nonlobotomized males under 60 with no evidence of brain damage in their clinical files. One hundred and one patients meeting these qualifications were originally chosen for the study but, because complete test data could not be obtained for seven Ss, the final *N* was 94. The mean age, education, and total months of previous neuropsychiatric hospitalization(s) for these men were 41.7 years ($SD = 8.2$), 10.8 years ($SD = 3.3$), and 118.8 months ($SD = 89.6$), respectively.

Measures taken on Ss were Becker's² 10-item revision of the Elgin Prognostic Scale, the Phillips Premorbid Scale, the Ullmann-Giovannoni PR questionnaire, the Barron *Es* Scale, months of previous psychiatric hospitalization, a quantified marital status measure (2 = married, 1 = separated or divorced, 0 = never married), education, and age.

Phillips Premorbid Scale and Elgin Prognostic Scale (10-item form) ratings had previously been obtained on all Ss as part of another study (Watson & Logue, 1968). Each author independently rated every S on both scales and the Elgin and Phillips scores used in the present study were totals over the two judges.

Product-moment correlations between each of the 28 pairs of variables were calculated. Subsequently the matrix was factored by the principal-factor solution (Harman, 1960, Ch. 9) using squared multiple correlation communalities in the diagonal. Two factors were then rotated orthogonally using Kaiser's (1958) normal varimax method.

RESULTS AND DISCUSSION

The correlation matrix is displayed in Table 1 and the composition of the factors in Table 2. The nature of the factors is discussed below. In defining the factors, only those variables with loadings of .40 or greater are considered.

² W. C. Becker, Modified Elgin Prognostic Scale, 1959. (Mimeo available from W. C. Becker, University of Illinois.)

TABLE 1
INTERCORRELATIONS OF PROCESS-REACTIVE MEASURES

Scale	2	3	4	5	6	7	8
Phillips (1)	-.67**	.60**	.70**	.36**	.15	-.17	-.03
Marital status (2)		-.51**	-.43**	-.32**	-.17	.15	.04
Ullmann-Giovannoni (3)			.45**	.33**	.08	-.07	-.10
Elgin (4)				.45**	.22*	-.06	-.17
Length of hospitalization (5)					.61**	-.19	-.26*
Age (6)						-.07	-.13*
Ego Strength (7)							.26*
Education (8)							

* $p < .05$.

** $p < .01$.

The first factor consisted of social adjustment measures and was defined largely by marital status and the Phillips, Elgin, and Ullmann-Giovannoni scales. More than the other factor, it seems to reflect conventional process-reactive measures. It accounted for 64% of the common factor variance.

A relatively small loading on this factor also appeared for length of psychiatric hospitalization, suggesting that the latter is not a good operational definition of the term although it is related to the process-reactive dimension. This finding also supports the contention made elsewhere (Lorei, 1966) that a surprisingly large share of the variance in length of neuropsychiatric hospitalization is attributable to nonpatient variables such as staff attitudes.

The second factor, which accounted for 36% of the common factor variance, consists essentially of age and length of hospitalization, although small loadings also appeared for education, *Es*, and Elgin scores. None of the more common process-reactive measures were highly

loaded on it and it appears to be of no particular value as a process-reactive measure.

It is particularly interesting to note that *Es* reflects quite a different dimension from that defined by marital status and the Phillips, Elgin, and Ullmann-Giovannoni scales. Even though the name "Ego Strength" suggests the scale as an appropriate process-reactive definition and it has been used as such in the past, the present data do not recommend it for that purpose.

The data suggest that marital status and the Phillips, Elgin, and Ullmann-Giovannoni scales are probably appropriate process-reactive definitions, but that neither *Es* nor length of previous hospitalization, age, or education would serve as an appropriate measure. Since the Elgin and Phillips scores used were composites based on the ratings of two psychologists, a comparison of the loadings recorded in Table 2 cannot be used as a basis for selecting a single process-reactive measure for future use. However, since the interjudge reliability of Phillips and Elgin scales is questionable (Watson & Logue, 1968), the data now available suggest that marital status and the Ullmann-Giovannoni scale may be the most appropriate operational definitions available, at least for male patients.

TABLE 2
ROTATED VARIMAX ORTHOGONAL MATRIX
OF FACTOR LOADINGS

Scale	<i>M</i>	<i>SD</i>	Factor	
			I	II
Phillips	44.5	14.5	.88	-.14
Marital status	40	.68	-.69	.14
Ullmann-Giovannoni	13.5	4.4	.66	-.12
Elgin	59.9	13.4	.67	-.30
Hospitalization	118.8	89.6	.31	-.74
Age	41.7	8.2	.07	-.65
Ego Strength	40.8	8.0	-.11	.22
Education	10.8	3.3	-.04	.34

REFERENCES

- BECKER, W. C. The relation of severity of thinking disorder to the process-reactive concept of schizophrenia. Unpublished doctoral dissertation, Stanford University, 1955.
- BECKER, W. C. A genetic approach to the interpretation and evaluation of the process-reactive distinction in schizophrenia. *Journal of Abnormal and Social Psychology*, 1956, 53, 229-236.
- HARMAN, H. H. *Modern factor analysis*. Chicago: University of Chicago Press, 1960.
- HERRON, W. G. Abstract ability in the process-reactive classification of schizophrenia. *Journal of General Psychology*, 1962, 67, 147-154.

- JOHNSON, M. H., & HOLMES, D. S. An attempt to develop a process-reactive scale for the MMPI. *Journal of Clinical Psychology*, 1967, 23, 191.
- KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, 23, 187-200.
- KANTOR, R. E., & HERRON, W. G. *Reactive and process schizophrenia*. Palo Alto, Calif.: Science & Behavior Books, 1966.
- KANTOR, R. E., WALLNER, J. M., & WINDER, C. L. Process and reactive schizophrenia. *Journal of Consulting Psychology*, 1953, 17, 157-162.
- LOREX, T. W. *Psychiatric Program Evaluation Staff studies of release decision making*. Washington: Veterans Administration Hospital, 1966. (Mimeo)
- PHILLIPS, L. Case history data and prognosis in schizophrenia. *Journal of Nervous and Mental Disease*, 1953, 117, 515-525.
- ULLMANN, L. P., & GIOVANNONI, J. M. The development of a self-report measure of the process-reactive continuum. *Journal of Nervous and Mental Disease*, 1964, 138, 38-41.
- WATSON, C. G., & LOGUE, P. E. A note on the interjudge reliability of Phillips and Elgin Scale ratings. *Journal of Clinical Psychology*, 1968, 24, 64-66.
- WITTMAN, P. *Scale for measuring prognosis in schizophrenic patients*. Elgin State Hospital Papers, 1941, 4, 20-33.

(Received February 19, 1968)

Journal of Consulting and Clinical Psychology
1969, Vol. 33, No. 1, 122-124

INTERNAL VERSUS EXTERNAL CONTROL AND ETHNIC GROUP MEMBERSHIP:

A CROSS-CULTURAL COMPARISON¹

THEODORE TIN-YEE HSIEH, JOHN SHYBUT, AND ERWIN J. LOTSOFF

Northern Illinois University

To investigate a relationship between Internal versus External control and ethnic group membership three groups of high school students (Chinese, American-born Chinese, and Anglo-American) were compared in their responses to the Internal-External Control Scale. A hypothesis that a belief in Internal versus External control is significantly related to ethnic group membership was confirmed.

Internal-External (I-E) control has been described as a major personality variable (Rotter, 1966) and its relationship with ethnic group membership has been reported (Battle & Rotter, 1963; Graves, 1961). In contrast to previous studies which compared different ethnic groups within the United States culture, the present study compares Chinese, American-born Chinese, and Anglo-American Ss in their responses to the Internal-External Control Scale.

The I-E may be defined as: "the extent to which an individual feels that his own actions determine what happens to him or what happens to him is determined by fate, luck, chance, or powerful forces outside his control [Rotter, Seeman, & Liverant, 1962]." A similar distinction in terms of the locus of control is also observed in the description of certain culture-specific personalities. Thus, the "individual-centered" American personality is associated with a culture that

emphasizes the uniqueness, independence, and self-reliance of each individual (Hsu, 1963a). It, among other things, places a high value on personal output of energy for solving all problems; pragmatic ingenuity; individualism, that is, self-reliance and status achieved through one's own efforts; power or ability to influence or control others, things, animals, and forces of nature; and efficiency (Gillin, 1955). Life experiences appear to be largely a consequence of one's actions.

In contrast, the "situation-centered" Chinese personality is associated with a culture where kinship and emphasis on status quo are stressed. The individual in this culture tends to view his life as being relatively fixed. Luck, chance, and fate are taken for granted in life, which is considered to be full of ambiguity, complexity, and unpredictability (Hsu, 1953, 1963b). Life situations may be viewed as being largely determined by circumstances outside one's control (Eberhard, 1965).

Consequently, it may be assumed that the internalized cultural orientation, other things being

¹This study is based on an M.A. thesis submitted by the first author to the Graduate School, Northern Illinois University.

equal, will be the discriminating variable in one's preference for the items stressing a belief in Internal versus External control. A hypothesis that, in comparison with American Ss, the Chinese will show higher preference for items emphasizing *external* control, was tested in this study.

METHOD

The original Internal-External (I-E) scale (Rotter, 1966) was translated from English to Chinese by the first author. The scale includes 29 forced-choice items such as: "No matter how hard you try some people just don't like you vs. People who can't get others to like them don't understand how to get along with others." In translating the scale all items were translated literally. Popular Chinese sayings whenever applicable were not used to avoid social desirability. In general, the principle of translation was comparable to that of Gough and Sandhu (1964), who translated the socialization scale of California Personality Inventory for cross-validation in India. Recently the I-E scale has been translated to Swedish and used with Swedish Ss.²

To test the Chinese translation, three judges who knew both English and Chinese were asked to translate independently the Chinese version into English. Their versions were very similar to each other and to the original English version. Subsequently, the Chinese version of the I-E scale was administered to 16 Chinese students at the local university. Three weeks later, the 16 students were given the English version of the scale. A highly significant correlation ($r = .83$) between the two sets of scores supports the reliability of the translated version of the I-E scale.

Subjects

The Ss for this study consisted of three groups of high school students. Group I included 131 male and 108 female Anglo-American high school seniors from a northern Illinois community near Chicago.

Group II was composed of 38 male and 42 female American-born Chinese high school seniors attending various public schools in Chicago. At least one parent of each S was born in China.

The Ss in Group III were drawn from three upper grades at Pui-Ching Middle School in Hong Kong. The school is a Baptist institution with an enrollment of over 5,000. Generally, it is attended by students from families of the middle or middle-upper classes, the majority of whom are not church-related or church-going. Of the 343 Ss in this group, 241 were males.

The Ss were closely matched for age (17.5, 17.2, 17.4 were the means for Groups I, II, and III respectively) and checked for socioeconomic background in terms of family income and father's education.

² M. Seeman, unpublished manuscript, 1964. Cited by J. B. Rotter (1966).

Procedure

The English version of the I-E scale was administered to Groups I and II while the Chinese version was administered to Group III with the following instructions:

Try to answer these items as *carefully* and *honestly* as you can. Do not spend too much time on any one item, but be sure to find an answer for every choice. Find the number of the item on the answer sheet and blacken the letter *a* or *b*—whichever you choose as the *more true statement*.

In some instances you may discover that you believe both statements or that you believe neither one to be true. In such cases, be sure to *select the one you more strongly believe to be true* as far as you are concerned. Also, try to answer each item independently when making your choice; do *not* be influenced by your previous choices.

REMEMBER!

Make the choice which *you personally believe to be more true*.

The forms were then identified and additional variables such as sex, grade, father's education, and family income were also collected.

RESULTS

The mean Anglo-American score was 8.58, $SD = 3.89$. The mean Hong-Kong Chinese score was 12.07, $SD = 3.96$. The highly significant difference between these two groups ($t = 10.57$, $df = 508$; $p < .001$) strongly supports the hypothesis (the scale is scored such that a higher score = more external).

The fact that the American-born Chinese Ss (mean score 9.79, $SD = 3.07$) are significantly more internally oriented than the Hong-Kong Chinese but less than the Anglo-American Ss further supports the acculturation hypothesis. The differences between these groups attained the .01 level of significance.

To check for possible interaction effect between I-E, ethnic group membership, father's education and family income, an analysis of variance was performed. The results of the analysis showed that only the among groups difference was significant ($F = 45.3$, $df = 2$, $p < .005$).

DISCUSSION

The findings of this study are relatively clear-cut. They support the postulated relationship between Internal versus External control and ethnic group membership. Heeding the warning of previous investigators, that socioeconomic variables are also related to Internal versus External control and should not be overlooked, two demographic indexes were used in this study to

account for socioeconomic level: father's education and family income. The main differences on the Internal versus External control dimension among groups remained when the groups were controlled for socioeconomic level. No attempt was made to control for intellectual functioning since most of the studies reported in the literature indicate very low and nonsignificant correlations (Rotter, 1966).

It appears that a cultural orientation may be closely linked with a personal belief in Internal versus External control. Individuals raised in a culture that values self-reliant individualism, pragmatic ingenuity, and personal output of energy are likely to be more internally oriented than individuals from a culture that tends to emphasize a different set of values.

REFERENCES

- BATTLE, E., & ROTTER, J. B. Children's feelings of personal control as related to social class and ethnic group. *Journal of Personality*, 1963, 31, 482-490.
- EBERHARD, W. Fate in traditional Chinese life. Paper presented at the meeting of the Association for Asian Studies. San Francisco, April 4, 1965.

- GILLIN, J. National and regional cultural values in the United States. *Social Forces*, 1955, 34, 107-113.
- GOUGH, H. G., & SANDHU, H. S. Validation of the CPI Socialization Scale in India. *Journal of Abnormal and Social Psychology*, 1964, 68, 544-547.
- GRAVES, T. D. Time perspective and deferred graduation pattern in a triethnic community. Research report No. 5, Tri-Ethnic Research project, University of Colorado, Institute of Behavioral Sciences, 1961.
- Hsu, F. L. K. *Americans and Chinese: Two ways of life*. New York: Akerlard-Schuman, 1953.
- Hsu, F. L. K. American core value and national character. In F. L. K. Hsu (Ed.), *Psychological anthropology*. Homewood, Ill.: Dorsey Press, 1963. (a)
- Hsu, F. L. K. Kinship and ways of life: An exploration. In F. L. K. Hsu (Ed.), *Psychological anthropology*. Homewood, Ill.: Dorsey Press, 1963. (b)
- ROTTER, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 1966, 80(1, Whole No. 609).
- ROTTER, J. B., SEEMAN, M., & LIVERANT, S. Internal versus external control of reinforcement: a major variable in behavior theory. In N. F. Washburne (Ed.), *Decisions, values and groups*. Vol. 2. London: Pergamon Press, 1962.

(Received April 10, 1968)

TREATMENT OF A PHOBIA BY PARTIAL SELF-DESENSITIZATION:

A CASE STUDY

MAX W. RARDIN

University of Wyoming

The treatment of a disabling case of hemophobia by partial self-desensitization is reported. The client completed the last nine items of a 16-item fear hierarchy by herself at home. Relaxation was self-induced and she presented the desensitization monologue to herself. At a 1-year follow-up she functioned as a student nurse on an obstetrics ward with no indication of recurrence of the phobia. The use of subvocal cues in controlling phobic behavior is discussed. The importance of the discrimination process in phobic reactions is suggested.

This paper presents the treatment of an incapacitating fear of blood and personal injury by a variation of Wolpe's (1966) desensitization technique. The purpose of the report is not to add to the list of exotic fears treated by desensitization but, rather, to illustrate a variation in technique, which to the author's knowledge has not appeared in the literature.

HISTORY AND CLINICAL DATA

The client was an 18-year-old single female in the first year of a nursing program and had been a student in an introductory psychology class taught by the author. At the encouragement of her nursing instructors, she contacted the author about the possibility of controlling her fear of blood. An interview indicated that she had a limited phobia. She reported no dissatisfaction in other areas of her life, was achieving above average grades, and had good relations with family, peers, and faculty.

The following history of the problem was gathered during the first scheduled session. The client indicated that she had been fearful of blood and generally squeamish for several years but her fears had not been a serious concern until she entered nursing—a career goal for her since childhood.

Her reaction to blood and possible physical injury varied from moderate discomfort to dizziness and nausea depending on the topic and circumstances. The immediate concern was her reaction to the films shown in nursing classes which vividly depicted various medical conditions. On a number of occasions, she had to put her head down or leave the room. She felt she would faint or vomit if she continued to observe the film. This reaction was interfering with her perform-

ance in the classes in which the films were shown, and the nursing faculty was beginning to question her suitability for the profession. Both S and her instructors felt it was imperative that she gain control over her reactions before the start of clinical classes in the hospital or she would not be able to continue in the program.

TREATMENT PROCEDURE

During the latter part of the first session, S was given a detailed description of desensitization as it would apply to her situation. The next three sessions were devoted to training in relaxation and construction of a fear hierarchy. The client was capable of deep relaxation and apparently experienced vivid imagery. The 16 items in the fear hierarchy involved increasing amounts of blood due to injury, surgery, and childbirth.

Except for changes in the placement of some items in the hierarchy, desensitization had proceeded routinely through seven items during the next four sessions when the school term ended. The seven items completed were: a scraped elbow; a torn hangnail; squeezing out one drop of blood; a cut in the sole of the foot; compound fracture of the leg; needle in skin for a stitch; and a gash in arm with flowing blood.

The remaining nine items which had not been dealt with were: bleeding from nose and mouth due to internal injury; a sucking chest wound; seeing a blood sample drawn; blood foaming from mouth; waters breaking for childbirth; head emerging and effect on mother; blood flowing after birth; delivery of placenta; and stitching after delivery. At this point, in order to continue treatment, the client was faced with the alternative of commuting 100 miles or attempting an experimental self-desensitization procedure suggested by the author.

The experimental procedures were explained as follows. The relaxation technique and the general style of the monologue accompanying the presen-

¹Requests for reprints should be sent to Max Rardin, Department of Psychology, University of Wyoming, Laramie, Wyoming 82070.

tation of each item was reviewed. She was reminded that the goal was to imagine increasingly vivid and personally relevant scenes. The importance of not proceeding if anxiety interrupted her relaxation was stressed. She was given general examples of monologue to accompany the rest of the items. On request, the client was able to produce additional appropriate examples. The possibility of arranging an observation of childbirth at a hospital was discussed. After further assurance that contact with the author would be available if there was difficulty, the client decided to attempt this program.

It was agreed she would work on the hierarchy each night when in bed under the following conditions. She was instructed to retire in her usual fashion, induce relaxation, and to begin with the highest item already completed. Work with the items was to be limited to approximately 1 hour. She was told to omit the procedure if she were sick, unusually tired, or disturbed by other matters. She later reported engaging in self-desensitization 5 or 6 nights a week for 6 weeks.

At the end of the 6 weeks she contacted the author and reported being able to imagine comfortably all of the items on the list and having visited the hospital maternity ward. She was late for the delivery but did see the cord being cut and the delivery of the placenta. At that point she felt mildly faint and left. She requested that she be allowed to use smelling salts at her next birth observation since her dizziness was not accompanied by nausea. Because she attributed this faintness more to excitement than anxiety, she was given permission to use smelling salts with the condition that she not force herself to observe if she felt highly anxious. She observed her next delivery successfully. The last session occurred after her return to school and was primarily a review of events to complete the case history. She reported having observed a complete delivery, successfully taken blood samples, and having her own blood sample taken. One year later she was a student nurse on an obstetrics ward fully assisting in deliveries to the point of

dabbing blood between vaginal stitches. Surgical repair after delivery had been the highest item on her hierarchy.

DISCUSSION

The client's progress with self-desensitization was apparently routine with no report of difficulty other than the incident which resulted in the request for smelling salts. At the conclusion of the self-desensitization procedure in a tape recording describing her thoughts on the experience, she revealed a technique she had improvised for maintaining control of her fear when approaching an anxiety-arousing situation. As she neared the situation, she would subvocally repeat to herself phrases abstracted from the desensitization monologue. The phrase "it's not me" was most often used. Her use of this phrase as an aide to discrimination appears similar to that of "it's just a dream" applied in the treatment of a recurrent dream reported by Greer and Silverman (1967). Both studies suggest that perhaps the discrimination process is crucial in controlling phobic reactions.

The need to use such phrases might raise a question about the degree of desensitization. Since she was doing so much of the work on her own, it may well be that desensitization was continuing at the time she reported her impressions. She later reported a decline in the use of the phrases.

This case suggests that self-desensitization may be a useful variation from the usual technique for clinical application. In research it offers a method by which the contribution of the therapist to desensitization might be more precisely defined.

REFERENCES

- GREER, J. H., & SILVERMAN, I. Treatment of a recurrent nightmare by behavior modification procedures: A case study. *Journal of Abnormal Psychology*, 1967, 72, 188-190.
- WOLFE, J., & LAZARUS, A. *Behavior therapy techniques*. New York: Pergamon Press, 1966.

(Received March 11, 1968)

BRIEF REPORTS

GROUP AND INDIVIDUAL ACCURACY IN PERSON PERCEPTION¹

RAYMOND E. FANCHER

University of Rochester

The purpose of this research was to compare the accuracy in person perception of judges making their decisions individually and in group discussion settings. The judges (*J*s) employed in this study were 48 male undergraduates from the University of Rochester, all of whom had had several courses in psychology. Each *J* participated once in an individual session and once in a group session with two other *J*s.

The instruments used to measure accuracy in person perception were two shortened versions of "programmed cases" that had been employed by the author in earlier research (Fancher, 1966). In a programmed case true events from the life history of a real person are presented in multiple-choice format, with alternative events from other persons' lives. A *J*'s task is to select what he thinks are the true events; after each of his choices he is informed as to which of the events was in fact the true one, so as he proceeds through a case he has increasing true information on which to base his later choices.

In the individual sessions, *J*s worked by themselves and recorded their choices. Their accuracy scores were simply the numbers of true events that they chose. In the group sessions, each *J* announced his choices to the group, and the group discussed each item until it reached unanimous agreement as to the most likely choice. The announced choices of each individual *J* as well

as the unanimous group choices were recorded and assessed for accuracy.

Each programmed case entailed 12 choices, and each *J* worked on one case individually and the other in a group. The research design was counter-balanced so that each case was used half the time in each condition, and so that half of the *J*s worked first individually and the other half worked first in a group.

The mean accuracy score for individuals working by themselves was 5.44 ($SD = 1.37$), while the accuracy scores for unanimous group choices and the announced choices of individual *J*s in the groups were 5.19 ($SD = 1.46$) and 4.83 ($SD = 1.34$), respectively. The only statistically significant difference was between the first score and the last ($t = 2.35, p < .03$).

Thus the results indicated that group discussions did not lead to increased accuracy, and that individuals announcing their choices in a group were actually less accurate in their judgments than when they worked alone. The group sessions were also highly inefficient, requiring an average of 50 minutes time on the part of the *J*s. The individual sessions averaged only 20 minutes in duration.

These results seem to suggest that further research attention might profitably be directed toward the investigation of certain clinical institutions like the "diagnostic council," which rest upon the assumption that interpersonal decisions can be made most accurately and efficiently in group settings.

REFERENCE

- FANCHER, R. E. Explicit personality theories and accuracy in person perception. *Journal of Personality*, 1966, 34, 252-261.

(Received June 13, 1968)

¹Reprints or an extended report of this study may be obtained without charge from Raymond E. Fancher or for a fee from the American Society for Information Science. Order NAPS Document 00212 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001; remitting \$1.00 for microfiche or \$3.00 for photocopies.

SOCIAL DESIRABILITY IN CHILDREN:

AN EXTENSION AND REPLICATION¹

EDWARD B. KLEIN

Yale University

LAURENCE J. GOULD

City College of the City University of New York

AND MICHEL COREY

Veterans Administration Hospital, West Haven, Connecticut

Crandall, Crandall, and Katkovsky (1965) have reported on the development of a children's social desirability scale (CSD). They found that social desirability (SD) scores were higher for younger than older children, duller than brighter Ss, girls than boys, and Negro than white children. Ordinal position, size of family, and social class were not associated with CSD. The authors also raised questions about the generality of the SD response and the underlying rationale for SD scales.

In light of the above questions, the present study sought to replicate the Crandall et al. findings and to further explore the meaning of SD by including sociometric data and ratings of children's behavior by nonsophisticated judges. The strategy called for a large sample of boys and girls drawn from as many settings as possible to test for generalizability and to obtain independent judgments of the children's behavior from naïve observers.

Method

The Ss for this study were 1008 boys and girls ranging in age from 7 to 14 years old. They were drawn primarily from a number of camps in the greater New Haven area during the summer of 1964.

The 48-item, true-false version of the CSD, a sociometric device (asking S to name the three children he likes and dislikes) and a background form were group-administered. At the same time counselors were asked to rate each S in terms of his friendliness, leadership, social awareness, and seeking approval from peers and adults.

¹ An extended report of this study may be obtained without charge from Edward B. Klein, Psychiatry Department, Yale University, 34 Park Street, New Haven, Connecticut 06519, or for a fee from the American Society for Information Science. Order Document 00211 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001; remitting \$1.00 for microfiche or \$3.00 for photocopies. Requests for reprint should be sent to Edward B. Klein at the above address.

Results

Differences between mean CSD scores at each age level by sex were tested by *t* tests. The results showed that girls score significantly higher than boys at every age level, except in the 12-year-old group. Additional *t* tests were carried out between age levels within each sex. In general, younger Ss of both sexes score higher than older Ss. This tendency is most marked when Ss are more than a year apart and younger than 11.

In keeping with Crandall et al., CSD score did not significantly relate to social class, birth order, and number of siblings. Consistent with the construct of SD, CSD scores were found to be significantly correlated with Ss' inability to voice a dislike of peers. Contrary to predictions CSD was not found to be related to independent ratings of counselors' perceptions of Ss.

Discussion

This study has demonstrated a replication of the Crandall et al. findings with respect to age and sex in a larger and more heterogeneous group of youngsters. Our findings also suggest that age 11 may be critical in that the CSD score appears to stabilize at this time. In addition, the significant negative correlation obtained between number of children rated as disliked and CSD occurs in both sexes primarily from age 11 onward. This finding would appear to provide some construct validity for the scale and suggests that the kinds of correlates associated with adult approval motivation begin to emerge in samples as young as 11 years old.

REFERENCE

- CRANDALL, V. C., CRANDALL, V. G., & KATKOVSKY, W.
A children's social desirability questionnaire. *Journal of Consulting Psychology*, 1965, 29, 27-36.

(Received June 20, 1968)

THE BAYLEY SCALES of INFANT DEVELOPMENT

Nancy Bayley

for babies
2 to 30
months
of age

Norms based on a stratified standardization sample controlled for age, sex, color, and education of parents.

Mental and Motor Scales yield separate scores:

Mental Development
Index (MDI)

Psychomotor Development
Index (PDI)

The child's interpersonal and affective behavior and the overall quality of his performance also are rated by the examiner on the accompanying Infant Behavior Record.



Photo by Marc & Evelyne Bernheim
Rapho-Guillumette Pictures, Inc.

Write for free descriptive leaflet.

Ψ The Psychological Corporation

312 East 45th Street, New York, New York 10017

**THE TOKEN ECONOMY:
A Motivational System for Therapy and Rehabilitation**

TEODORO AYLON and NATHAN H. AZRIN. The authors' pioneering work at Anna State Hospital (Illinois) results here in the detailed account of an effective new application of learning theory to the behavior problems of institutionalized patients. Representative of a departure from the current system of treatment in mental hospitals is the use of a wide variety of innovative reinforcers to strengthen desired behaviors, including original experiments that evaluate specific procedures. 288 pp., illus., paper, \$3.95

PSYCHOLOGICAL COUNSELING, Second Edition

EDWARD S. BORDIN, *The University of Michigan*. Up-to-date new edition includes an entirely new chapter on process problems in vocational counseling and coverage of recent developments in behavior therapy and theory and in research in personality and vocational choice. 472 pp., \$7.00

THE GOALS OF PSYCHOTHERAPY

Edited by ALVIN R. MAHRER, *Miami University*. Sixteen prominent psychiatrists, psychologists, and psychoanalysts join the editor in this assessment of modern psychotherapy. In discussing the goals of their profession, the contributors draw on the basic theories of personality, psychopathology, and human nature. The volume establishes a new basis for structuring the field of psychotherapy, in terms of a three-family organization of psychotherapies, illustrated with samples from the text. 301 pp., \$6.00

THE CLINICAL STUDY OF SOCIAL BEHAVIOR

DONALD R. PETERSON, *University of Illinois*. Seeks to arrive at a clearer definition and more effective resolution of the issues confronting clinicians in their study of behavior. Examining basic conceptual and methodological structures, the author suggests a number of solutions to practical problems as more helpful than those presently available. 254 pp., illus., \$5.50

THE HALFWAY HOUSE MOVEMENT: A Search for Sanity

HAROLD L. RAUSH and CHARLOTTE L. RAUSH, both of *The University of Michigan*. The first survey of halfway houses for the mentally disturbed in the United States. The volume discusses, among other topics, the role of the halfway house in meeting the problem of transition from the hospital to the community, preventive aspects of mental health, and the value of theories of social and psychological intervention. 247 pp., \$5.50

THE WORK OF THE COUNSELOR, Third Edition

LEONA E. TYLER, *University of Oregon*. A substantial revision—approximately fifty percent—of a highly lucid, inclusive and successful introductory text in counseling, this volume emphasizes the importance of creating and maintaining the kind of human relationship upon which the counseling process rests. The third edition includes a new classification system implemented for distinguishing between kinds of cases and used as a framework for organizing the material. March, 289 pp., \$6.50 (tent.)

Scheduled for 1969:

**A SOCIAL HISTORY OF HELPING SERVICES:
Clinic, Court, School and Community**

MURRAY LEVINE and ADELIN LEVINE, both of the *State University of New York at Buffalo*. The first detailed history relevant to the community mental health movement, this collaborative work integrates psychological, sociological, and historical data and concepts. 275 pp., \$6.50 (tent.)



APPLETON-CENTURY-CROFTS

EDUCATIONAL DIVISION

Meredith Corporation

440 PARK AVENUE SOUTH, NEW YORK 10016

EUROPE

A flexible travel program has been developed in connection with the XIXth International Congress of Psychology in London, July 27/August 2, 1969. Several groups will leave New York to Europe, the Mediterranean and the Near East -- utilizing the low group fares of Pan Am, TWA, BOAC and other major airlines. Round trip jet fares start at \$230 (New York/London/New York.) Write for brochure to: Division 'P', Travel Wholesalers International, 1707 L Street, N.W., Washington, D.C. 20036.

NEW IN 1969

DEVELOPMENTAL PSYCHOLOGY

Editor: Boyd R. McCandless

Developmental Psychology will contain articles which represent the broad range of growth and development and their major associated variables. Chronological age as well as sex, socioeconomic status and efforts of physical growth variables (including time of maturation and body build) are all considered relevant development variables.

Although the adolescent and the aging population appear to deserve particular attention at this time, *Developmental Psychology* will indeed cover the entire span. Cross-species articles and articles concerning developmental research with retardates will also be included.

The articles in *Developmental Psychology* will not be limited to any one approach such as "experiments versus studies" or "cognition versus motivation".

Bimonthly. One volume per year.

Subscription Price: \$10.00 per year

**AMERICAN PSYCHOLOGICAL ASSOCIATION
1200 Seventeenth Street, N. W.
Washington, D. C. 20036**

Journals of the AMERICAN PSYCHOLOGICAL ASSOCIATION

Subscription Rates for Calendar Year 1969

AMERICAN PSYCHOLOGIST. Official papers of the Association and articles on Psychology. Monthly. The 1969 volume is Vol. 24. Subscription: \$10.00 (foreign \$10.50). Single copy \$1.00.

CONTEMPORARY PSYCHOLOGY. Critical reviews of books, films, and research material in the field of psychology. Monthly. The 1969 volume is Vol. 14. Subscription: \$10.00 (foreign \$10.50). Single copy \$1.00.

JOURNAL OF ABNORMAL PSYCHOLOGY. Theoretical and research conditions in behavior pathology. Bimonthly; first issue appears in February. The 1969 volume is Vol. 74. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00. (Continues volume sequence of the former Journal of Abnormal and Social Psychology.)

JOURNAL OF APPLIED PSYCHOLOGY. Applications of psychology in business and industry. Bimonthly; first issue appears in February. The 1969 volume is Vol. 53. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF COMPARATIVE AND PHYSIOLOGICAL PSYCHOLOGY. Original contribution in the field of comparative and physiological psychology. Beginning in 1969: Monthly. Three volumes per year, four issues in each volume. The 1969 volumes are Vols. 67, 68 and 69. Subscription: \$30.00 for 3 vols. (foreign \$30.50). Single copy \$3.00.

JOURNAL OF COUNSELING PSYCHOLOGY. A primary publication medium for research in counseling theory and practice. Bimonthly; first issue appears in January. The 1969 volume is Vol. 16. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF CONSULTING AND CLINICAL PSYCHOLOGY. Research in clinical psychology: psychological diagnoses, psychotherapy, personality, psychopathology. Bimonthly; first issue appears in February. The 1969 volume is Vol. 33. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF EDUCATIONAL PSYCHOLOGY. Studies of learning and teaching: measurement of psychological development, psychology of school subjects, methods of instruction, school adjustment. Bimonthly; first issue appears in February. The 1969 volume is Vol. 60. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY. Original contributions of an experimental character. Monthly. Beginning 1969: Four volumes per year, three issues in each volume. The 1969 volumes are Vols. 79, 80, 81 and 82. Subscription: \$40.00 for 4 vols. (foreign \$40.50). Single copy \$4.00.

JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY. Theoretical and research papers on personality dynamics, group process, and the psychological aspects of social structure. Monthly. Three volumes per year, four issues in each volume. The 1969 volumes are Vols. 11, 12 and 13. Subscription: \$30.00 for 3 vols. (foreign \$30.50). Single copy \$3.00.

PSYCHOLOGICAL ABSTRACTS. Noncritical abstracts of the world's literature in psychology and related subjects. Monthly. Approximately 20,000 abstracts published annually. The 1969 volume is Vol. 43. Subscription: \$40.00 (foreign \$40.50). Single copy \$4.00.

PSYCHOLOGICAL BULLETIN. Evaluative reviews of research literature, discussions of research methodology in psychology. Monthly. Two volumes per year, six issues in each volume. The 1969 volumes are Vols. 71 and 72. Subscription: \$20.00 for 2 vols. (foreign \$20.50). Single copy \$2.00.

PSYCHOLOGICAL REVIEW. Original contributions of a theoretical nature. Bimonthly; first issue appears in January. The 1969 is Vol. 76. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

NEW IN 1969 DEVELOPMENTAL PSYCHOLOGY. Studies on growth and development and their associated variables. Bimonthly; first issue appears in January. The 1969 volume is Vol. 1. Subscription \$10.00 (foreign \$10.50). Single copy \$2.00.

Discounts: AGENTS and BOOK DEALERS receive a 10% discount on all orders. There is no club arrangement. There are no quantity, cash, library, or other special discounts. *Payment in U.S. funds.*

★ Subscriptions are available on a January through December basis only ★

AMERICAN PSYCHOLOGICAL ASSOCIATION
Subscription Department
1200 Seventeenth Street, N.W.
Washington D. C. 20036

Consider this important text

INTERNATIONAL OPPORTUNITIES FOR ADVANCED TRAINING AND RESEARCH IN PSYCHOLOGY

A summary of current opportunities for advanced training, research, and international exchange in psychology in more than 80 countries. Information about each country is presented in concise, topical form.

- General description of the educational system—primary, secondary, and university
- Advanced Education and University Research in Psychology—Organization of University Studies and Psychological Research at the Universities
- Training and Research in Psychology outside the University
- Recognition of Psychologists
- National Organizations of Psychologists
- Bibliographic Sources—Sources of more detailed information about the universities
- International Exchange—Receiving Foreign Visitors and Visiting Other Countries

395 pages

Price: \$3.50

AMERICAN PSYCHOLOGICAL ASSOCIATION

1200 Seventeenth Street, N.W.

Washington, D. C. 20036

Please send.....copy(ies) of International Opportunities at \$3.50 per copy.

Name:.....

Address:.....

State:.....Zip:.....

☐ Remittance enclosed

☐ Bill me

JOURNAL OF CONSULTING and CLINICAL PSYCHOLOGY

April 1969

Vol. 33, No. 2

Contents

ARTICLES

- ✓ Childhood Parental Relationships of Homosexual Men: *Ray B. Evans*..... 129
- Childhood Parental Relationships and the Establishment of Gender Roles of Homosexuals: *Ralph H. Gundlach*..... 136
- Parental Relations and Male Homosexuality in Patient and Nonpatient Samples: *Evelyn Hooker*..... 140
- Validity of Clinical Judgments Based on Human Figure Drawings: *Zev William Wanderer*..... 143
- ✓ DAP: Back against the Wall: *Emanuel F. Hammer*..... 151
- ✓ Creativity, Awareness, and Liking: *Russell Eisenman*..... 157
- Repression-Sensitization: Its Relation to Adjustment and Seeking Psychotherapy among College Students: *Mark H. Thelen*..... 161
- Sexual Aspects of TAT Administration: *Joseph Masling and Sandra Harris*..... 166
- Anality and Time: *Tupper F. Pettit*..... 170
- ✓ Validity of Informants' Ratings of the Behavior and Symptoms of Psychiatric Patients: *Norris D. Vestre and Robert Zimmermann*..... 175
- ✓ Maternal Personality Correlates of Sons' Creativity: *George Domino*..... 180
- Identifying Neurologically Impaired Children through a Test of Auditory Perception: *David A. Sabatino*..... 184
- Effects of Rater Level of Functioning and Experience upon the Discrimination of Facilitative Conditions: *John R. Cannon and Robert R. Carkhuff*..... 189
- Reactions of a Psychotherapy Group to Ambiguous Circumstances Surrounding the Death of a Group Member: *Donald D. Kirtley and Joseph M. Sacks*..... 195
- Perceptual Recognition in Chronic and Acute Schizophrenics: *Eugene Ebner and Barry Ritzler*..... 200
- Psychotherapy Process as a Function of the Time Segment Sampled: *Norman J. Karl and Norman Abeles*..... 207
- Self-Disclosure and Interpersonal Functioning: *Charles F. Halverson, Jr., and Roy E. Shore*..... 213
- ✓ Effects of Sex, Illness, and Hospitalization on Daydreaming: *Ann Pytkowicz Streissguth, Nathaniel N. Wagner, and John C. Weschler*..... 218
- Use of Speech in the Inhibitory Control of Motoric Behavior of Schizophrenic Patients: *Beatrice A. Ashem*..... 226
- Self-Concept of the Married Psychiatric Patient and His Mate's Perception of Him: *Martin Harrow, David A. Fox, and Thomas Detre*..... 235
- Effects of Anxiety on Timed and Untimed Intelligence Tests: Another Look: *Larry W. Morris and Robert M. Liebert*..... 240
- Preference for Complexity in Acute Schizophrenics: *John M. Neale and Rue L. Cromwell*..... 245
- Patients' Goals and the Manifestation of Psychopathology: *Alan F. Fontana and Theodore Gessner*..... 247
- Relationship of Selected Psychosocial Variables to Prognostic Judgments: *James E. Cowden and Asher R. Pacht*..... 254
- BRIEF REPORTS
- Comparative Study of the Wechsler Preschool and Primary Scale of Intelligence and the Stanford-Binet Intelligence Scale, Form L-M, among Culturally Deprived Children: *Allan Barclay and Allan C. Yater*..... 257
- Revised Suicide Potential Scale: *R. W. Miskimins and Lowell T. Wilson*..... 258
- LIST OF MANUSCRIPTS..... 225, 234, 239, 253

Brief Reports

The *Journal of Consulting and Clinical Psychology* will accept Brief Reports of research studies in clinical psychology for early publication without expense to the author. The procedure is intended to permit the publication of soundly designed studies of specialized interest or limited importance which cannot now be accepted because of lack of space. Several pages in each issue will be devoted to Brief Reports, published in the order of their receipt without respect to the dates of receipt of the regular articles. Most Brief Reports appear in the first or second issue to go to press following their final acceptance.

An author who wishes to submit a Brief Report:

1. Sends the Brief Report, limited to one printed page and prepared according to the specifications given below.
2. Also sends to the Editor a full report of the research study, in sufficient detail to give a clear account of its background, procedure, results, and conclusions, which will be filed with the National Auxiliary Publications Service (NAPS) to insure indefinite availability.
3. Prepares at least 100 mimeographed copies of the full report, which the author will send without charge to all who request it as long as the supply lasts.
4. Agrees not to submit the full report to another journal of general circulation.

Specifications

Brief Report. The Brief Report should give a clear, condensed summary of the procedure of the study and as full an account of the results as space permits.

To insure that the Brief Report will be no longer than one printed page, its typescript, in-

cluding all matter except the title and the author's lines, must not exceed 85 lines averaging 42 characters and spaces in length. Set the typewriter margins for short lines of 42 characters, which are 3.5 inches long in elite typing, and 4.2 inches long in pica.

The manuscript of the Brief Report must be *double spaced* throughout. Except for its short lines, it follows the standard style of the 1967 revision of the *APA Publication Manual*. Headings, tables, and references are avoided or, if essential, must be counted in the 85 lines. Each Brief Report must be accompanied by a footnote in the style below, which is typed on a separate sheet and *not* counted in the 85-line quota:

¹ An extended report of this study may be obtained without charge from John Doe (giving the author's full name and address) or for a fee from the American Society for Information Science. Order NAPS Document _____ from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001; remitting \$_____ for microfiche or \$_____ for photocopies.

Extended report. Because the extended report is intended for photoduplication, and is not copy to be sent to a printer, its style should differ in several ways from that of other manuscripts: (a) The extended report should be typed with single spacing for economy in duplication. (b) Tables and figures should be placed adjacent to the text which refers to them. A caption should be typed below each figure. (c) Footnotes should be typed at the bottom of the page on which reference is made to them. In other respects, the full report is prepared in the style specified by the *Publication Manual*.

Style of Manuscripts. Manuscripts must be prepared in the style described in the *Publication Manual of the American Psychological Association (1967 Revision)*, obtainable for \$1.50 from the Washington office of the APA.

Early Publication. Authors may secure the early publication of an acceptable article in the first issue to go to press after its acceptance by arranging for the article to be printed as extra pages, increasing the number of pages that subscribers receive. Information about charges for early publication may be obtained from the Managing Editor.

National Auxiliary Publications Service. To save printing costs, materials supplementary to articles may be deposited with the NAPS. To obtain copies, order by Document Number, as cited in the reference footnote, from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001. The indicated fee must be remitted in advance.

Abstracts. Manuscripts must be accompanied by an abstract of 100-120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or they may be obtained from the Editor or from the APA Central Office.

Copyrighted material. The following policies govern reprinting or other forms of reproduction of APA-copyrighted journal materials. Permission by the American Psychological Association must be requested for reprinting all tables, figures, and text of more than 500 words (total from one article) in length. There will be a charge of \$5.00 per page for such permission except when the requester is the author of the article to be reprinted, or the article is to be reproduced in a limited number of copies solely for instructional purposes.

Reprints. Authors receive 50 reprints gratis. Additional quantities may be ordered when returning proofs.

CHILDHOOD PARENTAL RELATIONSHIPS OF HOMOSEXUAL MEN¹

RAY B. EVANS²

Loma Linda University School of Medicine

A 27-item questionnaire adapted from Bieber et al. was administered to 43 homosexual and 142 heterosexual men in order to determine whether certain items which Bieber et al. had reported as differentiating homosexual and heterosexual males who were in psychotherapy would also differentiate similar groups who had never sought psychotherapy. The data were based on retrospective self-reports regarding childhood fears and activities, and interparental and parent-child relationships. In the Bieber study, each patient's psychoanalyst completed the questionnaire, so that their data were based on analysts' reconstructions of patients' early lives. The present results were remarkably similar to those of Bieber et al. in revealing more "negative" features in the backgrounds of the homosexuals. Nevertheless, the results were considered neither to support nor refute the Bieber conclusions regarding the etiological importance of parental relationships in homosexuality.

A major conclusion of a study comparing homosexual and heterosexual men who were all in psychoanalytic therapy was that parental roles are paramount in the etiology of homosexuality (Bieber, Dain, Dince, Drellich, Grand, Gundlach, Kremer, Rifkin, Wilbur, & Bieber, 1962). Those authors described the "classical" pattern as one where the mother is close-binding and intimate with her son and is dominant and minimizing toward her husband, who is a detached (particularly a hostile-detached) father to the son. They concluded that any son exposed to that parental combination will likely develop severe homosexual problems. The Bieber study was based on extensive questionnaires completed by the analysts for each patient; the patients themselves were not aware of the study. Two series of questions proved especially useful in differentiating the homosexual and heterosexual groups, a Developmental Six Score³ (concerning childhood fears and activities) and a

Twenty Questions Score (relating to interparental and parent-child relationships).

There is an obvious risk in generalizing findings from patients in psychotherapy to a non-patient population. The purpose of the present study was to determine whether questionnaire items adapted from Bieber et al. would differentiate samples of heterosexual and homosexual men who had never sought psychotherapy.

METHOD

Subjects

The sample consisted of 185 American-born, Caucasian men between the ages of 22 and 47, who had at least a high school education, had never sought psychotherapy, and were living in the Los Angeles metropolitan area. All Ss were volunteers in a study of cardiovascular disease, but only the 43 homosexuals knew that aspects of homosexuality were also being studied. The latter volunteered, as homosexuals, through the cooperation of a Los Angeles based organization; they did not constitute a representative group of homosexuals. The 142 "heterosexual" Ss volunteered for the cardiac study through a number of sources, and there was no opportunity to develop the kind of rapport needed to elicit information about their sexual preferences and behavior. For purposes of this study, it was assumed they were all heterosexual, though there may have been homosexuals among them, which would tend to attenuate group differences.

The homosexual men ranged in age from 22 to 46, with a mean of 33.8 years ($SD = 7.1$); the heterosexuals ranged from 25 to 47, with an average of 39.3 ($SD = 4.4$), and the difference was significant ($t = 4.84, p < .001$). They were reasonably similar to

¹ This study was supported in part by the California Foundation for Medical Research. Appreciation is extended to Jessie Marmorston for supplying the heterosexual Ss, and to One, Inc., for their cooperation in recruiting the homosexual Ss. Computational aid was furnished by the Loma Linda University Scientific Computation Facility, supported in part by National Institutes of Health Grant No. FR 00276.

² Requests for reprints should be sent to Ray B. Evans, Department of Psychiatry, Loma Linda University, Loma Linda, California 92354.

patients in the Bieber study, where the homosexuals averaged approximately 35 and the heterosexuals approximately 38 years of age.

In education, the homosexuals ranged from 12 to 19 years, with a mean of 14.4 ($SD = 2.2$); and the heterosexuals ranged from 12 to 20 years, with a mean of 15.1 ($SD = 2.1$), with the difference approaching statistical significance ($t = 1.87$, $p < .10$). Again, they were relatively similar to the Bieber patients, who averaged approximately 15 years of education.

There was an obvious and expected difference between the two groups in marital status. Among the heterosexuals, 5% were single, 87% married, and 8% separated or divorced; 86% of the homosexuals were single, 5% married, and 9% divorced. In the Bieber et al. study, 8% of the homosexuals and 51% of the heterosexuals were married.

As to sibling constellations, there were 12%, 35%, 14%, and 40%, respectively, only, oldest, middle, and youngest children among the homosexuals; comparable figures for the heterosexuals were 8%, 28%, 27%, and 37%. Those distributions were not significantly different. Bieber et al. reported 10% of their homosexuals and 22% of their heterosexuals were only children, which difference was significant at the .05 level.

Proportionately more of the homosexuals were employed in clerical work and in the arts, and fewer of them in other professions, management, and sales work. The difference in heterosexual-homosexual occupational distributions was significant ($\chi^2 = 42.88$, $df = 5$, $p < .001$). The occupational classifications of the present Ss and the Bieber patients were fairly comparable.

The homosexual volunteers rated their sexual experience on a 7-point scale adapted from Kinsey, Pomeroy, and Gebhard (1948), which ranged from entirely heterosexual to entirely homosexual. Of the 43 Ss, 58% described their experience as having been exclusively homosexual, 35% as predominantly homosexual with incidental heterosexual, and 7% as predominantly homosexual but more than incidental heterosexual experience. For their homosexual patients, Bieber et al. reported 68% as exclusively homosexual, 28% as having some heterosexual experience, and 4% as inactive, so that the proportion of exclusive homosexuals in the two studies was similar ($\chi^2 = .89$, $p = .50$).

The homosexual Ss also completed an 11-item questionnaire designed to determine their sexual identification. In their overall feelings, 40 Ss (93%) considered themselves moderately or strongly masculine, and responses to the other 10 items also indicated essentially masculine identification. Bieber et al. reported that approximately 2% of their homosexual patients were "markedly effeminate," so that seemingly Ss from the two studies were similar in this regard.

Procedure

Each S completed a 27-item questionnaire adapted from Bieber et al. so as to be as nearly comparable

as possible. The essential content of the questionnaire appears in Table 1. Included were the Developmental Six (Items 2-7) and Twenty Questions (Items 8-27) scores, and one additional question regarding physical make-up in childhood (Item 1), which had also differentiated the Bieber groups. Four possible choices were provided for each item, whereas the Bieber study used a yes-no dichotomy for all except three items. Following is an example from Bieber et al. completed by the analysts: "Was patient excessively fearful of physical injury in childhood? (yes/no)." The corresponding item modified for the present study, for completion by S himself, was: "During childhood, were you fearful of physical injury? (seldom/sometimes/often/always)."

Questionnaires were used in the analysis only where all 27 items had been answered, which eliminated 11 potential Ss, one homosexual and 10 heterosexual. Differences between groups were calculated by means of chi-square, with twofold classifications corrected for continuity.

RESULTS

The content of each questionnaire item is given in Table 1, together with the proportion of Ss responding in each category, and the significance of differences between the homosexual and heterosexual groups. Comparable figures from the Bieber study are also included, with the significance values calculated from figures in Appendix A (Bieber et al., 1962) using only Ss for whom definite response was available. Differences between the Bieber groups reached at least the .05 level of significance for 24 of the 27 items, and the other three approached significance. (Bieber et al. reported all items significant, and the discrepancy may be due to the method used here of ignoring "No Answer" and "Not Applicable" categories in the calculation of chi-square.)

In the present study, homosexual-heterosexual differences were significant at the .05 level or less for 24 items, one other approached significance, and for two items no difference was found. Thus, despite the very different method of collecting data, the non-patient status of the Ss, and (perhaps minor) differences due to geographical location, the results were remarkably similar to those reported by Bieber et al.

Specifically, in retrospect, the homosexuals more often described themselves as frail or clumsy as children and less often as athletic. More of them were fearful of physical in-

TABLE 1
QUESTIONNAIRE CONTENT AND ITEM RESPONSES

Questionnaire item	Bieber study				Present study			
	Response	Homo-sexual	Hetero-sexual	p	Response	Homo-sexual	Hetero-sexual	p
Physical make-up as a child	Frail	50	17		Frail	37	11	
	Clumsy	24	08		Clumsy	14	06	
	Athletic	13	33		Athletic	05	45	
	Well coordinated	13	42	.001	Coordinated	44	38	.001
Fearful of physical injury as a child	Yes	75	46		Seldom	23	49	
	No	25	54	.001	Sometimes	51	46	
					Often	19	04	
					Always	07	01	.001
Avoided physical fights	Yes	90	56		Always	56	12	
	No	10	44	.001	Often	30	35	
					Sometimes	14	46	
					Never	00	07	.001
Played with girls before adolescence	Yes	34	10		Never	09	03	
	No	66	90	.001	Sometimes	49	83	
					Often	40	14	
					Always	02	00	.001
"Lone wolf" in childhood	Yes	61	27		Never	12	38	
	No	39	73	.001	Sometimes	35	51	
					Often	42	11	
					Always	12	01	.001
Played competitive group games	Yes	17	64		Never	09	01	
	No	83	36	.001	Sometimes	65	15	
					Often	23	52	
					Very often	02	32	.001
Played baseball	Yes	16	64		Never	19	05	
	No	84	36	.001	Sometimes	70	29	
					Often	09	35	
					Very often	02	32	.001
Father and mother spent time together	Great deal	01	13		Great deal	16	28	
	Average	42	50		Considerable	53	39	
	Little	36	24		Little	26	23	
	Very little	21	13	.002	Very little	05	09	.23
Parents shared similar interests	Yes	20	38		Great many	21	30	
	No	80	62	.01	Several	37	32	
					Few	35	33	
					None	07	05	.70
Mother insisted on being center of son's attention	Yes	64	36		Never	30	18	
	No	36	64	.001	Seldom	37	63	
					Often	16	17	
					Always	16	01	.001
Mother "seductive" toward son as a child	Yes	57	34		Highly	07	00	
	No	43	66	.002	Moderately	07	03	
					Slightly	09	13	
					No	77	85	.02
Mother discouraged masculine attitudes/activities	Yes	39	17		Often	05	02	
	No	61	83	.002	Sometimes	21	07	
					Seldom	30	14	
					Never	44	77	.001
Mother encouraged feminine attitudes/activities	Yes	36	12		Never	53	87	
	No	64	88	.001	Seldom	21	11	
					Sometimes	21	02	
					Often	05	01	.001

TABLE 1—(Continued)

Questionnaire item	Bieber study				Present study			
	Response	Homo-sexual	Hetero-sexual	<i>p</i>	Response	Homo-sexual	Hetero-sexual	<i>p</i>
Mother considered puritanical	Yes	67	51	.05	Strongly	28	11	
	No	33	49		Moderately	33	35	
Mother's relationships with father/other men	Frigid	72	56	.04	Mildly	23	23	.04
	Not frigid	28	44		No	16	30	
Mother allied with son against father	Yes	63	40	.002	Frigid	12	00	.10
	No	37	60		Cold	26	23	
Mother openly preferred son to father	Yes	59	38	.005	Warm	63	77	.001
	No	41	62		Often	33	06	
Mother interfered with heterosexual activities	Yes	37	25	.08	Sometimes	21	18	.004
	No	63	75		Seldom	16	35	
'Son was mother's confidant	Yes	52	36	.03	Never	30	42	.004
	No	48	64		Always	12	01	
Son was father's favorite	Yes	08	29	.001	Often	14	06	.001
	No	92	71		Seldom	21	31	
Felt accepted by father	Yes	23	48	.001	Never	53	62	.005
	No	77	52		Often	12	00	
Son spent time with father	Great deal	03	03	.001	Sometimes	16	08	.006
	Average	12	39		Seldom	19	20	
Father encouraged masculine attitudes/activities	Little	37	31	.07	Never	53	71	.001
	Very little	48	27		Often	30	27	
Aware of hating father as a child	Yes	48	61	.002	Seldom	19	32	.001
	No	52	39		Sometimes	23	36	
Afraid father might physically harm him	Yes	57	43	.06	Often	28	05	.001
	No	43	57		Strongly	09	09	
Accepted father	Yes	21	51	.001	Moderately	16	40	.005
	No	79	49		Mildly	40	37	
Respected father	Yes	30	49	.01	No	35	14	.001
	No	70	51		Strongly	23	42	
					Moderately	35	42	
					Mildly	23	11	
					No	19	06	
					Great deal	02	08	
					Considerable	09	39	
					Little	53	32	
					Very little	35	21	
					Often	26	41	
					Sometimes	26	32	
					Seldom	23	21	
					Never	26	06	
					Never	28	59	
					Seldom	19	20	
					Sometimes	37	18	
					Often	16	03	
					Often	14	04	
					Sometimes	19	23	
					Seldom	30	13	
					Never	37	60	
					Strongly	26	51	
					Moderately	28	37	
					Mildly	33	09	
					No	14	03	
					Strongly	37	56	
					Moderately	21	32	
					Mildly	21	08	
					No	21	03	

Note.—Significance levels based on chi-square, with twofold classifications corrected for continuity. Decimals omitted.

jury, avoided physical fights, played with girls, and were loners who seldom played baseball and other competitive games. Their mothers more often were considered puritanical, cold toward men, insisted on being the center of the son's attention, made him her confidant, were "seductive" toward him, allied with him against the father, openly preferred him to the father, interfered with his heterosexual activities during adolescence, discouraged masculine attitudes, and encouraged feminine ones. The fathers of the homosexuals were retrospectively considered as less likely to encourage masculine attitudes and activities, and Ss spent little time with their fathers, were more often aware of hating him and afraid he might physically harm them, less often were the father's favorite, felt less accepted by him, and in turn less frequently accepted or respected the father. Unlike Bieber's patients, these homosexuals were no different from the heterosexuals in amount of time they estimated their parents spent together or in the interests shared by their parents.

In addition, a total score on the 27-item questionnaire was obtained for each S by weighting each item from 0 to 3 points, with the higher weighting at the "masculine" end, so there was a maximum possible score of 81. The scores of the homosexuals ranged from 9 to 64, with a mean of 42.9 ($SD = 11.6$); those of the heterosexuals from 36 to 77, with a mean of 57.3 ($SD = 9.2$). Though there was considerable overlap in scores for the two groups, the difference was highly significant ($t = 7.50, p < .001$).

DISCUSSION

The results could not be accounted for on the basis of sample characteristics other than sexual orientation; no relationship was found between age and questionnaire scores, and the same was true of marital status, occupational classification, and sibling constellation. The fact that the homosexuals knew homosexuality was being studied might have affected the results, but if there was any tendency to distort in the direction of "normal," it was not sufficient to obscure group differences.

As to preponderance of homosexual experience, no relationship was observed be-

tween Kinsey-type ratings (completed only by homosexual Ss) and questionnaire scores ($\chi^2 = 0.0$), perhaps because of the limited variation in proportion of homosexual experience. However, a product-moment correlation of .47 ($t = 3.41, p < .01$) was found between the 27-item questionnaire and the 11-item sexual identity questionnaire; the homosexuals with more "desirable" family backgrounds tended to consider themselves as more masculine.

It may be noteworthy that the present results were so similar to those obtained by the Bieber group despite a major difference in the level of observation. In the present study, the data were based on retrospective self-reports of how they now view their childhood, by Ss who had never been in psychotherapy. The Bieber data, on the other hand, were based on psychoanalysts' reconstructions of patients' early life circumstances, derived from impressions during psychotherapy. Arguments could be advanced for the superiority of one method over the other, and certainly both have limitations. The agreement in results could be interpreted as evidence of validity in both methods, or perhaps as an indication that the two methods are not essentially different.

The results strongly suggested poor parental relationships during childhood for the homosexual men, at least as seen in retrospect; however, the etiological significance of such relationships, or even the etiology of the relationships themselves, is another matter. Bieber et al. considered the chances high that any son exposed to the parental combination of maternal close-binding intimacy and paternal detachment-hostility will develop severe homosexual problems. Nevertheless, only 28% of their homosexual patients had such a parental combination, and the 11% of their control patients who had such parents did not become homosexual. Furthermore, Bieber et al. very much underemphasized one-third of the "triad," the son himself. They reported that "each parent had a specific type of relationship with the homosexual son which generally did not occur with other siblings," and that son was the "focal point for the most profound parental psychopathology." As to why a particular son

is singled out, Bieber et al. proposed that son is unconsciously identified by the mother with her own father or brothers, and the son thereby becomes the recipient of sexual feelings carried over from the mother's own early life. Similarly, the father transfers to that son his unresolved hostility and rivalry with his own father/brothers. The above is an oversimplified summary of the Bieber formulation, but it does not exaggerate the neglect of the son's contribution to the triadic relationship, beyond eliciting parental transference feelings.

The personalities and behavior of parents undoubtedly affect a child's personality, but some consideration must be given to the notion that the child's innate characteristics at least partially determine parental reactions and attitudes toward him. For instance, that the father of a homosexual son becomes detached and/or hostile because he does not understand or is disappointed in the son is just as tenable as that the son becomes homosexual because of the father's rejection. Similarly, that a mother may be more intimate with and bind her homosexual son more closely because of the kind of person he is, is just as reasonable as the idea that he becomes homosexual because she is too binding and intimate. Bieber et al. did question whether paternal rejection and hostility were stimulated out of feelings of disappointment and failure because of the son's homosexuality, but concluded that was not likely since only 17% of the fathers were reported to have been aware of the son's homosexuality. Surely most parental reactions crucially affecting the child's personality occur when the child is far too young to be labeled homosexual or heterosexual. The Bieber group also concluded that the father's attitudes were not traceable to the fact that the sons were inadequate and unattractive children, since the mothers did not find them so. That the mothers did not find these sons unattractive is no indication the fathers did not; the evidence suggests the fathers did find them unappealing.

Judging from experience with adult homosexual males, O'Connor (1964) also refuted the idea that lack of a good father relationship was a consequence rather than a cause of homosexuality, on the grounds that would

make it difficult to account for the many homosexuals whose fathers were physically absent. Bene (1965) rejected the notion that the lack of a positive father relationship might be due to the son repulsing the father rather than the father repulsing the son, and she cited O'Connor's reasoning. That homosexuality occurs in sons whose fathers are physically absent is irrelevant to the fact that when the father is physically present the relationship with the homosexual son is often a poor one. Furthermore, that homosexuality occurs in the absence of a father not only detracts from the etiological significance of a poor paternal relationship but in fact supports the importance of other causal factors (possibly such as innate physical/personality characteristics of the son).

Information was obtained relevant to another conclusion of Bieber and colleagues, who stated: "We have come to the conclusion that a constructive, supportive, warmly related father *precludes* the possibility of a homosexual son; he acts as a neutralizing, protective agent should the mother make seductive or close-binding attempts [p. 311]." The questionnaire responses gave no full and complete answer as to whether these fathers were constructive, supportive, and warmly related, but there was evidence that the father relationship of some homosexuals was as good as that of many heterosexuals. A score was calculated for each S based on Items 20-27 in Table 1, all of which concern the father-son relationship. With 0-3 points possible for each item, the total scores for the homosexuals ranged from 0 to 22, with a mean of 11.7; and scores for heterosexuals ranged from 3 to 23, with a mean of 16.5. While the difference was significant ($t = 5.95$, $p < .001$), 16% of the homosexuals scored above the heterosexual mean, and 16% of the heterosexuals scored below the homosexual mean. Therefore, it would seem that a moderately good relationship, at least as reflected in the above questionnaire items, does not preclude the appearance of homosexuality, even though it is well established that a poor father relationship is common among homosexual sons. The responses for the Bieber homosexuals on the corresponding questionnaire items suggested their father relationships

were poorer than those of the present homosexual Ss, which could merely reflect methodological differences, but more likely is related to the fact that the Bieber Ss had all sought psychotherapy, whereas none of the present Ss had done so.

In a similar fashion, two other questionnaire scores were calculated, one regarding mother-son relationships (Items 10-19) and the other pertaining to development (Items 1-7). With a maximum score of 30 on the 10 mother items, the homosexuals ranged from 4 to 29, with a mean of 18.8; the heterosexuals ranged from 13 to 29, mean 22.9. Although the difference was significant ($t = 5.39$, $p < .001$), 30% of the homosexuals scored above the heterosexual mean, and 12% of the heterosexuals scored below the homosexual mean. As to the seven developmental items, with a possible score of 21, the homosexuals ranged from 1 to 15, with a mean of 8.9; and the heterosexuals ranged from 6 to 19, mean of 14.3. That difference was most significant ($t = 10.39$, $p < .001$), and only 2% of the homosexuals exceeded the heterosexual mean, with 4% of the heterosexuals scoring below the homosexual mean. Of the three content areas, then, the developmental items clearly differentiated Ss best, with the

father and mother items similar in their differentiation. The childhood behavior reflected in some of the developmental items, of course, is not unaffected by parents, but the findings suggest the possibility of something more fundamental in homosexuality than a poor father relationship.

The results of the present study agreed closely with those obtained by Bieber et al. but they neither supported nor refuted the Bieber conclusions as to causal relationships. The complicated problem of the etiology of homosexuality probably could be more productively investigated with a prospective study.

REFERENCES

- BENZ, E. On the genesis of male homosexuality: An attempt at clarifying the role of the parents. *British Journal of Psychiatry*, 1965, 111, 803-813.
- BIEBER, I., DAIN, H. J., DENCE, P. R., DRELLICH, M. G., GRAND, H. G., GUNDLACH, R. H., KREMER, M. W., RIFEIN, A. H., WILBUR, C. B., & BIEBER, T. B. *Homosexuality: A psychoanalytic study*. New York: Basic Books, 1962.
- KINSEY, A., POMEROY, W. B., & GEBHARD, P. H. *Sexual behavior in the human male*. Philadelphia: Saunders, 1948.
- O'CONNOR, J. Aetiological factors in homosexuality as seen in Royal Air Force psychiatric practice. *British Journal of Psychiatry*, 1964, 110, 381-399.

(Received December 11, 1967)

EDITOR'S NOTE

The following two discussions by Ralph Gundlach and Evelyn Hooker were prepared at the request of the Editor.

CHILDHOOD PARENTAL RELATIONSHIPS AND THE ESTABLISHMENT OF GENDER ROLES OF HOMOSEXUALS

RALPH H. GUNDLACH¹

Postgraduate Center for Mental Health, New York, New York

Evans' questionnaire returns distinguishing "never-in-therapy" male homosexuals from comparisons confirm Bieber et al. findings. Evans' argument, that an association between questionnaire items regarding parent-to-child behaviors and later homosexuality is not causal, is rejected as not applicable. Studies by Money and Stoller with hermaphrodites and transsexuals establish the importance of parental belief about their child's sex in determining its sense of gender identity. Data regarding lesbians contrast with that for male homosexuals, and raise questions about multiple roads to adult sex-role character including: Does the partner's sex always have the same significance to S's self-concept?

It is with considerable satisfaction I learn that Evans' (1969) 27-item questionnaire—taken from the 450 items used by the Bieber team (Bieber, Dain, Dince, Drellich, Grand, Gundlach, Kremer, Wilbur, & Bieber, 1962)—answered by a sample of West Coast male homosexuals and heterosexuals never in psychoanalysis, confirm so solidly our findings on these items. We are grateful to Evans for testing our explicit generalization from a patient population to nonpatients.

In considering Evans' research I am reminded that there are many different and appropriate goals for research and that misunderstandings may occur between researchers because they do not grasp the others' context. I may very well fail to grasp Evans' through inability to get away from my own set. It seems to me that Evans presents his data with admirable objectivity, but I suspect he really did not expect these findings, for he veers off into a discussion which offers no accounting for his data, nor comes to grips with the Bieber et al. data. His arguments are not based on his research.

I see his argument as threefold: (a) he points out that the Bieber et al. findings cannot be taken to prove that parental relationships of the quality indicated have any etiological significance for homosexuality. One could just as reasonably assert, he says, that "the father of a son becomes detached and/or hostile because he does not understand or is disappointed in the son." Likewise, "the mother may be more intimate with and bind her homosexual son more closely because of the kind of a person he is." Although these statements might seem to hold logically for

the bare sets of answers to the 27 questions, Evans does not take account of the many cross-tabulations between related questions, and especially such interrelationships that touch differences in time as presented in Chapter VII, "Developmental aspects of the prehomosexual child" (Bieber et al., 1962). But Evans himself adds—undermining his argument it seems to me—"Surely most parental reactions critically affecting the child's personality occur when the child is far too young to be labeled homosexual or heterosexual."

(b) Evans argues that even though fathers of homosexuals (H) and controls (C), for instance, may be distinguished at the .001 level or higher on a collection of items, the two groups still overlap in part. He thus seems to require that the answers to questionnaire items about relations between parents and child as remembered and categorized in simplified general phrases, whether answered by a psychiatrist or the adult son, should be as firm as a causal connection; as solid as the copper wire between light switch and light bulb.

This position seems to indicate a misconception about the nature of the subject matter of the research and a confusion about the function and meaning of questionnaire data in studies such as these on homosexuality. Evans follows here, perhaps, the old medical model regarding the etiology of a disease, like mumps or yellow fever, and demands no less than the identification of one precise "cause." But the facts show that homosexuality is not a disease, or a syndrome, or a unitary trait; and certainly the questionnaire items are not the causal factors. Evans erects and attacks a straw man so far as he purports that Bieber et al. assert that "a poor father relationship" or a close-binding-intimate

¹Requests for reprints should be sent to Ralph H. Gundlach, 160 East 84th Street, Apt. 2K, New York, New York 10028.

mother "causes" homosexuality in a son. The whole tenor of that study was to delineate the multivariate factors that may be involved. Particularly impressive were our findings of a strong heterosexual interest in many of the H which were sharply inhibited by fears related to female genitalia; and evidence that many H seek pickups not for sheer sexual pleasure and gratification or love, but are driven by anxieties and are acting compulsively. Such neurotic fears are hardly instinctual, but do they also characterize H who never had psychotherapy? Unfortunately Evans did not replicate these most essential items.

(c) Evans' third argument is that the Bieber et al. study overlooked the son himself: one-third of the triad. He asserts "the child's innate characteristics at least partially determine parental reactions and attitudes toward him."

It is hard to understand Evans' objection that the individual personality was left out in the Bieber et al. study, since the patient was the focus of the entire study, triangulated from many vantage points. The study was an attempt to enrich, extend, and systematize clinical material about male homosexuals found in psychoanalytic treatment; the focus was always on the ground level of the particular person. The superposition of an extensive questionnaire was understood as a gross way of collecting disparate persons of family groups. The questionnaire answers and the many tables served to block out the main geography of patterns of male homosexuals and their family of origin patterns, just as a map serves to fix the relative sizes and some of the characteristics of ground-level reality. The statistical material was treated not as primary data but as material to be assessed in the light of clinical experience and to extend clinical insights. Thus, some eight classifications of mother-types and 10 groupings of father-types provided 51 different parental patterns; and these were further collected into four major groupings. Throughout, with every section, the full meaning was illuminated, the clinching material was displayed through some 35-40 cases. We found a remarkable number of conditions and circumstances that may result in homosexuality.

When Evans suggests that there is something within the nature of the homosexual which may be the determining factor ("possibly such as innate physical/personality characteristics of the son"), one could think he was unhappy with his own findings. But the evidence continually mounts against such an "innate" hypothesis. The study of genes, chromosomes, and endocrine secretions have all been largely discredited as determining homosexuality. Perloff (1965) sum-

marizes that "homosexuality is purely a psychological phenomenon, neither dependent on a hormonal pattern for its production nor amenable to change by endocrine substances [p. 68]."

What happens after the child is born is complicated by many factors; there are not only inner biological and emotional factors, parental and familial surroundings, social and cultural circumstances; but the various pressures and expectations shift as the child grows and hardens as he establishes his ways into his eventual adult character structure.

At birth the baby is identified as a boy or girl and the child is reared by the parents with that idea in mind, with whatever doubts, reservations, pleasures, or disappointments, and intentions they carry for the child. So by the age of 2 or 3 it knows that it is a male or a female; its core gender role is established. But at this time, according to Rabban (1950), the child has little of the sex-role characteristics, which are largely established after age 3. Still later come the pressures and anxieties about being acceptable as a masculine or feminine person.

One's sex and gender role are not always consonant, and some most interesting studies have been made of persons caught in this predicament. Money (1965), for instance, describes two cases that were hermaphrodites and biologically the same. But one was raised as a girl and the other as a boy and each developed appropriately and without difficulty into these opposite gender roles. Other more confusing cases arise where the external genitalia give the appearance of one sex—this is announced by the obstetrician at the child's birth—but these conceal internal genitalia carrying the primary characteristics of the opposite sex. The child is reared by parents who have a mistaken idea as to its sex; but when biological difficulties arise, usually at the approach of sexual maturity, the error in sex assignment is discovered. The earlier approach was to proclaim the biological truth, and expect the child to shift its gender role; but this turned out to be usually disastrous, courting psychotic breakdowns. Money (1965) and Stoller (1968) found it proved much better to let the established gender identification prevail, and refit the person's physical and secondary sex characteristics to the person's self-concept. To reverse the psychological structures of the self, incorporated around sexual identity and all the habits, attitudes, values, and ways of living this entails, is extremely hazardous.

Stoller made extensive and precise studies of some patients with biological abnormalities, and some without biological abnormalities but with

gender-role anomalies. In many of these cases not only the child but one or both parents were interviewed, several were in analysis, and the interactions especially of mother and child were discovered in detail. He concluded that differences in gender from biological sex are almost always produced by differences in infant rearing. Most interesting were those male cases he called transsexual. These were male babies that developed within the first year and a half a sense of female identity; and he found and studied adult transsexuals whom he assumed to have been reared as were these males. The adult transsexuals only wanted their male bodies to be changed so as to conform to their feminine outlook. These patients, he pointed out, had no touch of psychosis; they were quite sane. Further, this feminine identity was established not as an adaptation to severe pressures (see Rado, 1965), but under conditions of conflict-free pleasure, provided by an overindulgent, close-binding mother. Stoller points out how these transsexuals are sharply different from transvestites, fetishists who know they are male but periodically take a woman's role for special reasons; different from effeminate homosexuals who like to dress in women's clothing, usually to satirize and make fun of women. In all the child transsexuals he saw, there were some striking similarities. There was a close mother-son symbiosis, with the mothers tending to hold their babies against their own bodies for much too long a time; the mothers all had a special identifiable bisexuality and emptiness; they encouraged cross-dressing of their sons; and their husbands tended to be physically absent from the household, and were unconcerned with the cross-dressing and other behaviors of the son imitative of the mother. These transsexuals, he emphasized, never thought of themselves as homosexuals, since they had the complete sense of being women.

Probably none of the patients in the Bieber et al. study, and none of the women in the Gundlach and Riess (1968) study, had been faced with the biological and social problems just described. The males all accepted the fact they were males, although the H and many C had problems about potency and how masculine they were or were perceived to be; and how, by sexual aggression or other behavior they could promote their social reputation or enhance or protect their self-esteem. Likewise the women studied all knew themselves as females but their means of establishing or expressing their gender role is quite different in this society than for men. Girls learn femininity by absorption, without effort by imitation of mother (see Lynn, 1962).

They do not have to assert themselves to be women.

I would like to end by calling attention to two bits of evidence about female homosexuality, each of which raises interesting questions.

We find that the incidence of homosexuality in relation to birth order depends upon a number of factors (Gundlach & Riess, 1967). For instance, even with our limited sample of about 225 H and 230 C there were significantly (.05) more lesbians who were firstborn among families with only one or two children. However, of families with five or more children, 12 were firstborn but only three of these were H. Of these large families, 22 were fifth or later in rank, 17 H and five C; of the 48 who were in the first four ranks, 16 were H and 32 C ($p < .001$).

What makes this important is that quite different family dynamics must operate with the firstborn girls from small families than for later-born girls from large families. It might be that firstborns are pushed to produce, and feel they must prove themselves with accomplishments as boys must do, while the later-born girls in large families are perhaps largely neglected. (Often the last-born male is the one focused upon by mother, and becomes homosexual.) But furthermore, the question arises as to why it is that the firstborn girls in large families are not as equally prone to being lesbian as within smaller families? I cannot say, but it seems that something about the decision cannot occur until after it is established that there will be five or more siblings; and by that time the firstborn girl must be 10-15 years old!

The other exciting item comes from the study of the patterns of female homosexuality in a prison for women convicted of a felony. Ward and Kassebaum (1965) describe how the majority turned to homosexuality as a way of life: to seek closeness, warmth, and affection from another inmate as the way to survive the loneliness, degradation, and other psychological deprivations and injuries of incarceration. It seemed bound up with emotional needs. Yet 68% of the women were mothers; and only 5-15% were estimated to have been homosexuals "outside."

In men's prisons, on the contrary, males still feel they must demonstrate their masculinity by force: by fighting, sodomizing weaker ones, buying favors, and discharging debts with a sexual "favor." However, the persons who are forced to comply, or traffic in sex, are viewed with contempt by those rated as highly masculine. Yet these forceful leaders, who use the others sexually, do not consider themselves homosexuals.

This brings up the point that some men can have sexual relations only with men, or only with women, or with neither, or both, or with a fetish; and some are willing and able to engage in sexual activity with a great variety of objects and/or persons.

Most women, however, express the feelings that they cannot have sex, or do not wish to, without "love." But many find that they can love a man or a woman. Few are like the majority of male H in treatment, who seem to fear or even have a phobic reaction to heterosexual intercourse. So—what is the (varying) importance of the sex of the sexual partner, to the S's gender role and self-concept?

In the light of evidence of cultural determination of gender role and sexual practices, the possibility of an innate physical/personality characteristic determining homosexuality seems quite remote.

REFERENCES

BIEBER, I., DAIN, H. J., DANCE, P. R., DRELLICH, M. G., GRAND, H. G., GUNDLACH, R. H., KREMER, M. W., RIFKIN, A. H., WILBUR, C. B., & BIEBER, T. B. *Homosexuality: A psychoanalytic study*. New York: Basic Books, 1962.

- EVANS, R. B. Childhood parental relationships of homosexual men. *Journal of Consulting and Clinical Psychology*, 1969, 33, 129-135.
- GUNDLACH, R. H., & RIESS, B. F. Birth order and sex of siblings in a sample of lesbians and non-lesbians. *Psychological Reports*, 1967, 20, 61-62.
- GUNDLACH, R. H., & RIESS, B. F. Self and sexual identity in the female: A study of female homosexuals. In B. F. Riess (Ed.), *New directions in mental health*. New York: Grune & Stratton, 1968.
- LYNN, D. B. Sex-role and parental identification. *Child Development*, 1962, 33, 555-564.
- MONEY, J. *Sex research*. New York: Holt, Rinehart & Winston, 1965.
- PERLOFF, W. H. Hormones and homosexuality. In J. Marmor (Ed.), *Sexual inversion*. New York: Basic Books, 1965.
- RABBAN, M. Sex role identification in young children in two diverse social groups. *Genetic Psychology Monographs*, 1950, 34, 81-158.
- RADO, S. A critical examination of the concept of bisexuality. In J. Marmor (Ed.), *Sexual inversion*. New York: Basic Books, 1965.
- STOLLER, R. J. *Sex and gender*. New York: Science House, 1968.
- WARD, D. A., & KASSEBAUM, G. G. *Women's prison: Sex and social structure*. Chicago: Aldene, 1965.

(Received June 29, 1968)

PARENTAL RELATIONS AND MALE HOMOSEXUALITY IN PATIENT AND NONPATIENT SAMPLES¹

EVELYN HOOKER²

University of California, Los Angeles

A comparison is made of studies employing patient and nonpatient samples. Evans has made an important contribution to etiological studies of male homosexuality by replicating the Bieber questionnaire items with a nonpatient sample. Nonpatient samples are important because of the possibility of ruling out the contaminating factor of psychopathology other than homosexuality per se. Serious questions are raised about the role of parent-child disturbed relations as necessary and sufficient conditions leading to adult male homosexuality. Other etiological variables are briefly stated.

It can no longer be questioned that faulty, disturbed, or pathological parental relationships in early childhood are more commonly reported by male homosexual patients than by a comparable group of male heterosexuals. The studies of West (1959), Braaten and Darling (1965), Freund and Pinkava (1961), O'Connor (1964), and Bieber, Dain, Dince, Drelllich, Grand, Gundlach, Kremer, Rifkin, Wilbur, and Bieber (1962) are in agreement on this finding although they differ with respect to the specific critical dimensions of the parental relationships. In some studies, the content of the disturbed relationship is very general and lacks specific behavioral criteria. Thus, for example, West finds that homosexuals more frequently have "overintense" relations with their mothers and "unsatisfactory" relations with their fathers. In O'Connor's study, homosexuals are more attached to their mothers and do not "re-late normally" to their fathers. On the other hand, Bieber et al. very carefully delineate a substantial number of behavioral criteria for the close-binding-intimate (CBI) mother and the hostile, detached father as illustrated by the items selected for cross-validation in the Evans' (1969) study.

Evans' careful replication of the Bieber study with nonpatients who answer questions directly concerning parental relationships is a very important contribution to the further clarification of the role of parental relationships in homosexuality. The etiological role of parental relationships in producing homosexuality is an *inference* which cannot be justified from psychiatric sam-

ples alone, in part because of the contamination of homosexuality with psychopathology. Similar results with nonpatient samples tend to confirm that inference even though they are not conclusive because of the necessarily large number of other uncontrolled variables in a retrospective study.

Prior to Evans', five studies of nonpatient samples have appeared. As with the patient samples, no precise cumulative picture emerges because the investigators used different dimensions of parental relations. Thus, Westwood (1960) found indifferent or poor relations with the father and an overprotective or possessive mother in approximately 50% of his sample. Unfortunately, the lack of a control group makes it impossible to assess the significance of this finding. In the Schofield (1965) study, homosexual men in a nonpatient sample reported a higher incidence of poor relations with the father and of overprotective or overpossessive mothers than did a similar sample of heterosexual men. Bene (1965) found no evidence that homosexuals were more often attached to their mothers or overprotected and overindulged by them. They more often reported more hostility and less affection from and toward both mothers and fathers than did heterosexual men. In Apperson and McAdoo's study (1968) mothers of homosexuals were found to be less restrictive than were the mothers of heterosexuals, and the fathers were described as more cold and impatient. Thus, the picture of the mother which emerges from these studies is contradictory but that of the father is fairly consistent from one study to another. However, in Greenblatt's (1966) study *no* differences were found between homosexual and heterosexual men in ratings on a semantic differential of attributes of their mothers and fathers. Indeed, fathers of both groups were rated by them as good, generous, pleasant, dominant, and under-

¹ This investigation was supported by a Research Career award from the National Institute of Mental Health, United States Public Health Service.

² Requests for reprints should be sent to Evelyn Hooker, Department of Psychology, University of California, Los Angeles, Los Angeles, California 90024.

protective and mothers as good, pleasant, neither subordinate nor dominant, and neither overprotective nor underprotective. None of these studies permits an exact comparison with the investigations of patient samples because of differences in methodology or in the definition of the relevant dimensions.

Evans' study is therefore of special importance as a partial confirmation of the Bieber assumption about causal relation between parental relations in early childhood and adult homosexuality. In my view, Evans is overly cautious in his assertion that his findings neither confirm nor refute the etiological role of parent-child relations as one set of many variables influencing or causing homosexuality in adult life. Indeed, his study necessitates this generalization since it is a replication whereas it is unwarranted from the Bieber study alone. His caution and the lack of appropriate caution in the Bieber study are both unwarranted extremes.

The issue of contaminating psychopathology with homosexuality, unfortunately, is not settled by the Evans' study since he did not control for this variable. In fact, the amount of psychopathology other than homosexuality per se in the nonpatient samples thus far investigated is unknown except for Greenblatt's study: Theoretically, nonpatients should show less psychopathology only if it is assumed that greater psychopathology is the impetus to treatment. However, as was demonstrated in the Manhattan and Stirling County investigations, psychopathology may be widespread in an untreated population. Psychopathology is only one set of the complex pattern of variables partially influencing whether treatment is sought.

Greenblatt took this into account in the selection of his nonpatient groups by administering the MMPI as a measure of psychopathology. Both of the homosexual and the heterosexual samples were found to consist of two subgroups: one as "normal" as any group in the MMPI literature and the other showing some psychopathology. However, there was more psychopathology in the homosexual "nonnormal" subgroup than in the heterosexual one. Nevertheless, the total group showed much less psychopathology than is true of neurotic Ss generally. The highest score on any clinical scale for the homosexual group was on the Depression Scale and this mean was only 60.4. Thus, the fact that Greenblatt's data do not confirm the Bieber findings may be due to the comparative lack of psychopathology in the total homosexual nonpatient group. On the other hand, it may be due to differences in methodology.

Some supportive evidence for interpreting the results as due to the relation between parental relations and adult psychopathology rather than to homosexuality per se is given by Schofield's finding that the incidence of disturbed parental relations was greater in the patient than in the nonpatient homosexual samples. Evans' finding that relations with the father were somewhat better, and in fact were "moderately good" in some instances, as compared with Bieber's sample, lends some additional support to the argument. Apperson and McAdoo's findings that mothers of homosexuals were less restrictive than those of the control Ss could also be interpreted in this way. The evidence, however tentative, of the relation between disturbed parental relations and adult psychopathology makes it imperative that further studies include an assessment of psychopathology in the samples studied.

The negative evidence of the Greenblatt study (no difference between parental relations in the homosexual and the heterosexual groups) and the findings in all studies that disturbed parental relations characterize only a portion of the homosexual sample as well as a smaller proportion of the heterosexual sample raise serious questions about the etiological role of parental relations in the genesis of homosexuality. Furthermore, studies of schizophrenic patients have demonstrated that a very similar picture of parental relations characterize the early childhood (Lidz, Cornelson, Fleck, & Terry, 1957).

Thus it appears that disturbed parental relations are neither necessary nor sufficient conditions for homosexuality to emerge. If it should be the case that they are more commonly found in the background of severely disturbed patients than of those individuals who function relatively well, the necessity of looking at other etiological factors should be underlined.

Evans arrives at his position that innate factors may be important by evaluating the significance of the father's role. He argues that because homosexuality develops in sons whose fathers are absent as well as in those with moderately good relations with fathers, the etiological role of the father must be questioned and the contribution of the son's characteristics (perhaps innate) must be assessed. While I am in complete agreement with Evans about potential genetic or other etiological contributions to a homosexual object choice (even though there is no possibility of specifying what they are with the evidence currently available), the force of his argument about the father's role is weak if one considers the function of the father in producing or

preventing homosexuality. That is, Bieber et al. assume that the presence of a warm, strong, and supporting father who intervenes on behalf of the son counteracts the effects of a CBI mother. The absence of the father, as well as a poor father, may leave the son in the double bind of maternal sexual seductiveness-maternal sexual restriction.

Finally, there are four major theoretical issues concerning the etiology and determinants of persistent or predominant adult homosexuality. (a) Is the human organism psychosexually neutral at birth, so that learning processes determine homosexual object choice in adults, or are there inherent sexual predispositions, which selectively influence the effects of learning? (b) What is the nature and content of the learning processes by which homosexual object choice develops? Is the appropriate developmental model a deviant role or a personality system with intrapersonal traits, motives, and gender identifications incompatible with the social-sexual capabilities and self-other expectancies of adult relations with the opposite sex? Does positive conditioning of sexual responses to persons of the same sex, or negative conditioning to persons of the opposite sex, or a combination of both account for homosexuality? (c) Are particular periods in the developmental process, such as early childhood or adolescence, critical for homosexual object choice? (d) Are parent-child relationships in the nuclear family crucial in determining whether an individual becomes homosexual, or are peer relationships in childhood and adolescence, and deviant subcultures in adolescence and early adult life, of equal or possibly greater importance? These four issues are highly controversial and cannot be resolved by the research evidence currently available.

REFERENCES

- APPERSON, L. B., & McADOO, W. G., JR. Parental factors in the childhood of homosexuals. *Journal of Abnormal Psychology*, 1968, 73, 201-206.
- BIEBER, I., DAIN, H. J., DINCE, P. R., DRELICH, M. G., GRAND, H. G., GUNDLACH, R. H., KREMER, M. W., RIFKIN, A. H., WILBUR, C. B., & BIEBER, T. B. *Homosexuality: A psychoanalytic study*. New York: Basic Books, 1962.
- BENE, E. On the genesis of male homosexuality: An attempt at clarifying the role of the parents. *British Journal of Psychiatry*, 1965, 111, 803-813.
- BRAATEN, L. J., & DARLING, C. D. Overt and covert homosexual problems among male college students. *Genetic Psychology Monographs*, 1965, 71, 269-310.
- EVANS, R. B. Childhood parental relationships of homosexual men. *Journal of Consulting and Clinical Psychology*, 1969, 33, 129-135.
- FREUND, K., & PINEAVA, V. Homosexuality in man and its association with parental relationships. *Review of Czechoslovakian Medicine*, 1961, 7, 32-40.
- GREENBLATT, D. Semantic differential analysis of the "Triangular System" hypothesis in "adjusted" male homosexuals. Unpublished doctoral dissertation, University of California, Los Angeles, 1966.
- LIDZ, T., CORNELISON, A. R., FLECK, S., & TERRY, D. The intrafamilial environment of the schizophrenic patient. II. Marital schism and marital skew. *American Journal of Psychiatry*, 1957, 114, 241-248.
- O'CONNOR, P. J. Aetiological factors in homosexuality as seen in Royal Air Force psychiatric practice. *British Journal of Psychiatry*, 1964, 110, 381-391.
- SCHOFIELD, M. *Sociological aspects of homosexuality*. Boston: Little, Brown, 1965.
- WEST, D. J. Parental figures in the genesis of male homosexuality. *International Journal of Social Psychiatry*, 1959, 5, 85-97.
- WESTWOOD, G. A. *A minority report on the life of the male homosexual in Great Britain*. London: Longmans, Green, 1960.

(Received July 12, 1968)

VALIDITY OF CLINICAL JUDGMENTS BASED ON HUMAN FIGURE DRAWINGS¹

ZEV WILLIAM WANDERER²

Teachers College, Columbia University

The literature establishes that clinicians have routinely been making diagnostic judgments based on techniques of equivocal validity. In a study considering clinical as well as experimental specifications for a suitable methodology, it was found that Draw-A-Person experts are capable of identifying mental defectives far beyond chance expectations. Four remaining matched groups (schizophrenics, neurotics, homosexuals, and normals), however, were found not to be identifiable, even after the experts were permitted a second chance to make a correct diagnosis. The expertness of the judges, as predicted from the ranks they were accorded by their peers, was unrelated to their actual performance.

A major responsibility of the psychologist is to insure that the techniques that he uses actually do what they purport to do. As Meehl once put it, "Regardless of one's theory about personality and regardless of one's choice of data, whether Rorschach, MMPI, . . . intuition, table, equation, or rational hypotheses developed in a case conference—the honest clinician cannot avoid the question 'Am I doing better than I could do by flipping pennies?'" [Meehl, 1954, p. 136].³ The assurance of validity is particularly urgent in the area of psychodiagnosis. Clinicians often contend that they do not use their tests in a mechanical, additive way, but that their decisions are ultimately complex judgments.

Clinical judgments, however, are not exempt from scientific scrutiny. Hunt concluded that "we should consider the individual clinician as a clinical instrument, and study and evaluate his performance exactly as we study and evaluate a test [Hunt, 1946, p. 317]."

At a 1958 Society for Projective Techniques symposium, Kenneth Little summarized the situation regarding the validity of projective techniques: "It is somewhat embarrassing to have to say that the published evidence on projective techniques indicates that they have either zero, or at best, very low positive effective validity indices. Even in those studies with the most positive of results, correlations are of an order of magnitude which make predictions for the individual largely a waste of time [Little, 1959, p. 287]."

In 1949, Karen Machover standardized the administration and formalized the interpretation of a figure-drawing test which later became known as the Draw-A-Person (DAP) test. Although she saw S's verbal associations to his drawings as an auxiliary source of personality information, she relied heavily on universal symbolism and other psychoanalytic constructs in the absence of, or in addition to, those associations. The decision to study the DAP as the subject of the current investigation was based on three criteria: its widespread use (Sundberg, 1961), contradictory conclusions reported in reviews of the literature (Jones & Thomas, 1961, pp. 256-258; Lublin & Lublin, 1967; Swensen, 1957, pp. 460-461), and the allegation that "few of

¹Based on the author's dissertation at Teachers College, Columbia University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy. Partially supported by the Behavior Research Institute of California (Grant 69-002).

²Currently with the Center for Behavior Therapy, 415 North Camden Drive, Beverly Hills, California 90210. The author wishes to express appreciation to Edward Joseph Shoben, Jr., who supervised the entire project; to Laurance F. Shaffer and Stanley Lieberfreund, who provided guidance in the report of the research; and to Richard Lindeman and William Frank, who were especially helpful with the statistical aspects of the study. Special thanks are due Karen Machover and Solomon Machover for their long consultations and thoughtful suggestions during the formative stages of the study. Valuable editorial assistance was provided by Aviva Wanderer. A brief version of this paper was discussed at the 74th Annual Convention of the American Psychological Association, New York, September 1966. For reprints contact the author at the above address.

Machover's hypotheses have been explicitly tested by definitive studies [Swensen, 1957, p. 460]."

Quest for a Method

A validating use of statistics that makes no assumptions beyond those of the theory of probability is the method of correct matchings (Meehl, 1954, pp. 11-12). Dating back to Vernon's (1936) early graphological experiments, the matching method keeps the criterion constant and turns to individual differences in the dependent variable. In that way, it differs from classical experiments, modeled after those in the physical sciences which systematically vary the independent variable and record changes in the dependent variable. In a matching study which uses judges who base their judgments on test data, the judges are the Ss, and their judgments are the dependent variable. Judgments should be expected to vary in response to the differential stimuli represented by the productions of discrete criterion groups. Hence, the method is called "differential" (Andrews, 1948). Tables, prepared by Chapman (1934) and Dudek (1952), provide expected frequencies for such forced-choice matches.

The requirements set forth in the American Psychological Association's (1954) *Technical Recommendations* concerning the report of validity information imply, by analogy, procedures for the proper validation of a test. Consideration of these recommendations along with the findings of relevant studies (e.g., Bolotin, 1960; Chambers & Hamlin, 1957; Dana, 1962; Dawes, 1962; Dennis, 1960; Mogar, 1962; Mosteller & Bush, 1954; Secord, 1952; Wallon, 1959), criticism (e.g., Brown, 1952; Hamlin, 1954; Hammer, 1959; Schneider, 1950; Shneidman, 1959), unpublished manuscripts, and personal communications by Karen Machover and Solomon Machover, provided enriching sources for conceptualizing the method elected here. In sum, the empirical literature suggested that a concurrent validity study, applying the method of correct matchings of global judgments (Cronbach, 1950) to clinically homogeneous groups, would be appropriate. Researchers have cautioned that judges be experts, that samples be relevant, that demographic and sampling

variables be controlled, and that the task be clearly defined—not too simple, yet not unmanageably complex.

Hypotheses

1. Diagnostic judgments, made by experts using the DAP, agree, beyond chance expectations, with criterion statuses.

2. Differences in accuracy of DAP judgments vary between clinical groups beyond chance expectations.

3. The judges' diagnostic accuracies are positively related to their reputed expertness, defined by ranks accorded them by their peers.

PROCEDURE

Consultations were held with the Machovers in order to determine which five clinical groups should be most appropriate to a diagnostic investigation of the DAP. Among those suggested were a psychotic, a neurotic, a mental defective, and a homosexual group, in addition to a group of normals. Other recommendations resulted in the decisions that all Ss be American Caucasian adult males residing in the northern United States.

The Criterion Groups

In order to avoid contamination of the criteria by the variable being investigated, Ss were selected only if they were never given the DAP for diagnostic purposes (except for the mental defectives, who are almost routinely given the DAP in state institutions, but who are certified as mental defectives on the basis of their scores on standardized IQ tests). The Ss who passed all of the selection criteria were administered the DAP de novo in accord with the instructions and materials detailed in Machover's (1949) monograph.

The sample groups were carefully matched for age, ethnic status, socioeconomic class, marital status, and other relevant variables, except for the mental defectives, whose very condition precluded their being matched for intelligence and education.^a

The overall principle in the determination of the external criteria for the group labels was that the five clinical groups represent different life situations, recognized by commonly encountered diagnoses. Thus, the schizophrenics had been so diagnosed and committed to a state hospital for most of their adult lives; the neurotics, on the other hand, lived and worked in the community, but attended an outpatient psychiatric clinic on a regular, weekly basis.

^a Tables and appendices detailing this, and other information, are embodied in the longer dissertation upon which this report is based. Microfilm and Xerox copies of the dissertation may be ordered from University Microfilms, Ann Arbor, Michigan, Pub. No. 65-7406.

The mental defectives achieved IQ scores well into the defective range and had been legally certified and committed to a state school for mental defectives. The homosexuals lived and worked in the community, did not receive any form of psychological treatment, denied that such treatment was ever officially indicated to them, belonged to a national society for homosexuals, and in each instance classified themselves as practicing homosexuals. The normals who were finally selected lived and worked in the community, attended an evening college, denied having been seen for psychological treatment, and obtained "nonpsychiatric" scores on the Cornell Medical Index.⁴

The Judges

Starting with Karen Machover, reputed experts in the interpretation of the DAP were asked to name and rank at least five other experts whom they considered particularly competent in the diagnostic use of the DAP. Each of the nominees was then asked to do the same, etc., until a pool of over 50 names emerged. The experts were contacted in the order of their ranks, so that the 20 highest ranking cooperating experts were designated the judges for this study.⁵

Preparation of the Data

Five Ss were selected from each of the criterion groups. Each S drew a person and then a person of the sex opposite to that represented in the first drawing. Each S and his pair of drawings were assigned a code from a table of random numbers (Kendall & Smith, 1950). The design called for 100 clinical judgments to be made by a total of 20 judges. Each judge received one pair of drawings from each of five clinical groups. Each pair was thus seen by four different judges. Twenty sets of five DAPs each were prepared for distribution in the following manner: Five envelopes were labeled by clinical groups; four copies of each S's code number were placed in the appropriate envelope; one number was drawn from each of the five envelopes to make up a set until all 20 sets were completed.

⁴Thanks are due the following who provided accessibility to clinical groups: Sherman O. Schachter and the staff at the New Hope Guidance Center, Brooklyn; Norman Kilpatrick and the officers of the Mattachine Society; Sidney Robbins and the staff at Central Islip State Hospital, N. Y.; Issac Wolfson and Thomas McCulloch and the staff at Letchworth Village State School, N. Y.; Manny Sternlicht and his students at the Staten Island Community College.

⁵The author is grateful to the 20 DAP experts who served as judges: Helen J. Anderson, Lauretta Bender, Fred Brown, Hanna F. Faterson, Bernice M. Gurvich, Emanuel F. Hammer, Molly R. Harrower, Asya L. Kadis, Selma L. Landisberg, Solomon Lieberman, Karen Machover, Solomon Machover, Muriel F. Margolis, Hanna S. Marlens, Frank S. Puzzo, Joseph Richman, Joanna Steinberg, Meta Steiner, Rochelle Wexler, and Mildred Zadek.

The Task

Each of the 20 judges received a package containing five pairs of drawings representing one each of the five diagnostic groups. Included also was an enclosure describing the characteristics of the sample, and a letter defining the task: "Identify each coded pair of drawings as to the clinical group from which you feel it has been drawn. This is done by ranking the probability of each diagnosis for each pair." Attached to each pair of drawings was a "Judgment Form" which instructed the judge,

Please look at the attached pair of drawings by the same person and decide which of the five groups they most likely represent. Place the number "1" on the line next to that group in the list below, for the diagnostic group that seems most probable. Continue ranking the categories by placing a "2" on the line near the next likely possibility, and so on, so that the category which is least likely to be represented by these drawings will be marked with a "5."

The five categories were listed below the instructions with an adjacent space for the rank numbers. The judges were also asked,

Try to pinpoint in four statements the basis for your decision in each case, e.g., a scoring system, formal elements in the drawings, content, a detail, global impressions, a hunch, or your own unique system of evaluating drawings.

Finally, each judge was asked to complete an "Experience Questionnaire."

Auxiliary Experiment

For expository purposes, the study was replicated but with the substitution of cards with chance-determined "judgments." That is, the code numbers of the Ss were set up on a results tabulation chart similar to the one used for the actual DAP results. The E called out the code number of an S, and an assistant⁶ picked a card out of a box and read the "diagnosis" from the card. In order to simulate the forced-choice restriction of the first choice, the card was not replaced until the next "judge's" turn came. This was done for all 20 "judges," for two (rather than five) diagnostic rankings. These randomly drawn data were then compared with theoretical chance and with judgments by DAP experts in appropriate tables.

RESULTS

Chance Expectations for the Experimental Procedure

In the auxiliary chance-illustration replication, the random drawings produced zero to

⁶Thanks are due Louis Garzetta, who painstakingly kept this procedure random, yet simulated the judges' actual task.

three "successes" per "judge" out of a possible five, with an average of one. The randomly drawn "successes" were tabulated. The chi-square computed to determine the significance of any difference between correct matchings produced by the random drawings and those expected by theoretical chance was .815, which, with $df = 2$, is nonsignificant. Successes for each criterion group, by this empirical method for demonstrating the effects of chance, averaged close to the expected 20%. Similarly, when accurate matchings on a second run of the random drawings were combined with those obtained on the first random drawings trial, the resulting cumulative "hits" resembled frequencies theoretically expected by chance. Thus, the "successes" averaged two out of a possible five (since there were two chances to be correct), and "successes" for each criterion group averaged 40%. The chi-square, computed to determine the significance of any difference between the successes obtained on the random drawings and those expected by chance, was .143, which, $df = 2$, is nonsignificant. The experimental procedure, stripped of the "experts using DAP" variable, generated results essentially identical with those expected by chance alone.

Successes Achieved by DAP Experts

The average number of successes by DAP experts was twice the number arrived at by

TABLE 1

SIGNIFICANCE OF THE DIFFERENCES BETWEEN JUDGES' SUCCESSSES AND THOSE EXPECTED BY CHANCE

Number of "hits"	Number of judges		Chi-square
	Theoretical chance	DAP experts	
First choice			25.82*
2 or more ^a	5	15	
1	8	4	
0	7	1	22.98*
First and second			
3 or more ^a	7	17	
2	7	2	
0 or 1 ^a	7	1	

Note — All decimals rounded to nearest integer.

^a Where individual χ^2 cells yielded expected frequencies less than 5, adjacent cells were combined.

* $p < .01$; $\chi^2_{.01} = 13.82$, $df = 2$.

the random drawings; the DAP successes for the criterion groups averaged close to 40%. The chi-square, computed to determine the significance of the difference between the successes of the DAP experts and those obtained by random drawings, was 40.42, which, $df = 2$, is significant beyond the .001 level.

After the first and second correct matchings were combined, a chi-square comparison constructed to show the frequency of randomly drawn correct matches over two trials, revealed, again, that the DAP judges' overall performance differed from that produced by random drawings at a highly significant level.

Chi-squares were computed to determine the significance of the differences between the successes of the DAP experts and those expected by chance. Since the computation of chi-square requires that no cell have an expected frequency of less than five, those cells which had fewer theoretical frequencies were combined (Siegel, 1956, p. 109). The resulting frequencies are listed in Table 1. For example, although only five judges would be expected to achieve two or more successes on their first-ranked diagnoses by chance alone, as many as 15 DAP experts scored that many successes. Similarly, although only seven judges would be expected to achieve three or more successes over their first and second choices combined (by chance alone), as many as 17 DAP experts scored that many successes. The highly significant chi-squares, when viewed superficially, seem to indicate that the demonstrated ability of the experts to discriminate between the five clinical groups, on the basis of the DAP, was not due to chance. Further analysis of the data reveals, however, that the mental defectives were detected almost perfectly when the judges used up their first ranks, and with 100% accuracy when the judges' first two diagnostic opportunities were combined.

Variances between Judges and between Clinical Groups

For diagnoses ranked first, one judge out of 20 scored a perfect five "hits" out of five possibilities, and one judge scored no "hits." Analyses of the variances between judges indicates that the differences are probably due to chance.

A wide range of successes by DAP experts for the five clinical groups was revealed. Analyses of the variances between groups indicated that the difference in the susceptibility of the various groups to be identified by means of the DAP is statistically highly significant.

The Contributions of the Mental Defective Group

The highly significant variance between groups raises a question: What is the contribution of the mental defective group to the finding that experts can identify clinical groups by using the DAP alone? This question was answered by dropping the mental defective group from the findings and testing the remaining data for significance. The "Total" columns of the accuracy tables were adjusted to show the distribution of the successes of the DAP experts in identifying only four of the original five clinical groups: the schizophrenics, neurotics, homosexuals, and normals. The nonsignificant chi-squares on Table 2 strongly suggest that the initially observed ability of the experts to discriminate between the four clinical groups, on the basis of the DAP, was probably due to the obviousness of the drawings by the mental defectives—especially when the experts had two opportunities to be correct.

Consideration of the Possibility of a Misleading Group

The statistical significance, then, of the original finding that experts can identify clinical groups by using the DAP vanished once the mental defective group was omitted. The fact that the schizophrenic group was misdiagnosed more often than any other group raised a different question: If the schizophrenic group were also omitted from the data, would the vanished significance reappear? The "Total" column of the accuracy table was adjusted to show the distribution of successes of the DAP experts in identifying just three of the original five clinical groups. The resulting nonsignificant chi-square indicated that the observed ability of the experts to discriminate the three clinical groups on the basis of the DAP (such as the correct identification of 11 out of 20 homosexuals) was still within the bounds of chance expectations. The

TABLE 2

JUDGES' DAP SUCCESSSES COMPARED WITH CHANCE
SUCCESSSES AFTER OMITTING THE MENTAL
DEFECTIVE GROUP

Number of "hits"	Number of judges		Chi- square
	Theoretical chance	DAP experts	
First choice			1.32 ^a
2 or more	5.83	7	
1	6.70	8	
0	7.50	5	
First and second			3.42 ^a
3 or more	6.42	6	
2	7.40	11	
0 or 1	6.18	3	

^a $ns; \chi^2_{.05} = 5.99, df = 2$.

schizophrenic group, therefore, was not sufficiently misleading to determine the chance-like performance of the experts when the mental defective group was omitted.

Distribution of Incorrect Judgments

A table was constructed to determine whether the misdiagnosed Ss were misdiagnosed consistently, and how they were identified by the judges. Table 3 indicates that the outstanding group was the mental defectives, which had only one misdiagnosis. Incorrect judgments for the rest of the groups were fairly scattered with no significantly consistent pattern.

Other Findings

1. Overall interjudge agreement was less than half as great as it would have been had

TABLE 3

DISTRIBUTION OF INCORRECT JUDGMENTS

Criterion status	Frequency of judgments					False negatives
	Schizo- phrenic	Neu- rotic	Men- tally defec- tive	Homo- sexual	Nor- mal	
Schizophrenic	—	7	0	6	4	85%
Neurotic	5	—	0	3	8	80%
Mentally defective	1	0	—	0	0	3%
Homosexual	6	0	1	—	2	45%
Normal	5	9	0	0	—	70%
False positives	85%	80%	5%	45%	70%	

all the judges agreed about all Ss. The greatest contribution to interjudge agreement was the mental defective group, about which there was 90% agreement.

2. The ranks for expertness that the judges received from their peers (in the early nomination procedure) were juxtaposed with the ranks the judges achieved by virtue of the accuracy of their clinical judgments. The resulting Spearman rank correlation coefficient, corrected for ties, was $-.36$, which is nonsignificant ($r_{.95} = .377$).

3. Attempts to analyze the judgmental process (judges' statements in support of their decisions), including the procedure used by Chambers (1954, pp. 17-18, 23-24, 33-34, 56-58), revealed that stylistic differences between DAP judges were the same as between Ss for any one judge. No differentiating pattern emerged.

4. Analysis of the "culture" of DAP experts revealed that women outnumbered men by better than two to one. There were five times as many experts who were primarily engaged in clinical work than in teaching and research; about two-thirds of the experts worked in institutional settings, and although 76% relied "heavily" on the DAP, only 32% regarded it as a "primary tool."

DISCUSSION

Adequacy of the Method

The differential findings—that the judgments were consistently valid for one criterion group and consistently invalid for the remaining groups—imply that the procedure contained no elements which would tend to favor or disfavor the object of the investigation. Comparison of the findings with those of Chambers' (1954) study of the Rorschach, which provided the foundation for our design, suggests that the DAP, not the method, was responsible for our results. As in the current study, Chambers found highly significant overall positive results: Her mental defectives were identified 90% of the time, far more frequently than were the four other groups. Omission of the mental defective group, however, hardly affected her findings; the resulting chi-square was 16.3, leaving the positive results unchanged and still significant beyond

the .001 level. Thus, application of the same basic design which achieved positive results with the Rorschach in the Chambers' study, achieved negative results with the DAP in our study. The negative results must be attributed to the DAP, and not the experimental design.

Five methodological innovations were added here to the basic correct-matchings design:

1. The Ss were administered the DAP *de novo* in order to avoid contamination of the criterion by the predictor.

2. A method was used for determining the relative contributions of individual criterion groups to the total results.

3. By requesting the judges to rank their diagnoses, the judges were freed from the restriction of having to make an absolute decision on a protocol about which they may have lacked confidence because of what they regarded insufficient or conflicting evidence.

4. A procedure for the identification of expert judges proved successful. The nomination and ranking procedure provided an additional source of information for determining the relationship between the judges' reputations among their peers and their relative actual performances.

5. The definition of criterion groups by life statuses, rather than by psychiatric diagnoses or psychological test results alone, provided an operational method of avoiding the hazards of unreliability or nonfactuality of diagnoses *per se*.

Why Then is the DAP Popular?

The studies which report positive findings for the DAP are primarily case studies. There is little doubt that there are instances when individuals produce sensationally revealing drawings. As Swensen (1957, p. 461) pointed out, such cases are more likely to be remembered than are the many which are not revealing. Furthermore, such cases are more likely to be submitted for publication than are the less eloquent ones.

Another conjecture about the cause for the positive regard which so many clinicians have for the DAP was advanced by Shaffer.¹ Prior to and during the administration of the

¹L. F. Shaffer, personal communication, 1958.

DAP, interview material and behavioral clues are typically observed by the clinician; upon interpreting the drawings that the client has produced, the clinician may attribute the interview and other extratest knowledge, which he has about S, to the drawings, "seeing" in them what he knew already. "Blind" studies eliminate this hazard.

Little has proposed an explanation for the failure of projective techniques to prove valid: the matter of situation generalization. The test tests the organism, he claimed, not the astronomical number of alternative stimulus situations in which the organism functions and about which the clinician attempts to predict. By means of generalization, it is possible to reduce the number of stimulus categories sharply. "We are able to say that one man's candy is another man's brandy and call both consummatory responses, or oral behavior, etc. This inevitably leads to a loss in precision in the prediction of specific behavior so we find ourselves anticipating alcoholism when obesity in fact occurs [Little, 1959, p. 288]."

Hammer has suggested that it is fallacious to conclude that drawings do not reflect psychopathology simply because associations between drawing signs and criterion groups have not been demonstrated.

All of these "schizophrenic signs" are relatively infrequent in projective drawings of the human figure, but where they do occur, they occur in the drawings of schizophrenics. . . . To test these hypotheses adequately, only instances where the sign does occur should be included. For example, to wait to accumulate twenty subjects' drawings and then compare the incidence of schizophrenia in the subjects who submitted these drawings, would be the only way to assess fairly the validity of the sign [Hammer, 1959, p. 30].

Essentially, Hammer is simply proposing to ignore false negatives in DAP predictions. A high incidence of true positives would mean that some drawings reflect some blatant characteristics of the Ss who produced them, but one could still wonder about how much confidence could be placed in an instrument which, used alone, would send a large proportion of schizophrenics merrily on their way, since they would be identified as anything but schizophrenics.

Perhaps, the most basic question to investigate is: What causes clinical psychologists to believe in and use instruments which repeatedly fail the test of diagnostic validity? Could the occasional, eloquent drawing provide the clinician with partial reinforcement, producing greater resistance to extinction? An experiment investigating Shaffer's⁸ hypothesis that beliefs of psychologists direct their reinforcing acts could shed much light and may prove convincing even to the committed.

⁸ L. F. Shaffer, personal communication, 1959.

REFERENCES

- AMERICAN PSYCHOLOGICAL ASSOCIATION. *Technical recommendations for psychological tests and diagnostic techniques*. Washington, D. C.: Author, 1954.
- ANDREWS, T. G. *Methods of psychology*. New York: Wiley, 1948.
- BOLOTIN, M. The use of human figure drawings in evaluating children and adolescents of special educational and cultural background; an examination of the effectiveness of current diagnostic criteria with the Draw-A-Person Test as applied to Puerto Rican children and adolescents. *Dissertation Abstracts*, 1960, 20, 4030. (Abstract)
- BROWN, F. House-Tree-Person and human figure drawings. In D. Bower & L. E. Abt (Eds.), *Progress in clinical psychology*. New York: Grune & Stratton, 1952.
- CHAMBERS, G. S. *An investigation of the validity of judgments based on "blind" Rorschach records*. (Doctoral dissertation, University of Pittsburg) Ann Arbor, Mich.: University Microfilms, 1954. No. 9964.
- CHAMBERS, G. S., & HAMLIN, R. M. The validity of judgments based on "blind" Rorschach records. *Journal of Consulting Psychology*, 1957, 21, 105-109.
- CHAPMAN, D. W. The statistics of the method of correct matchings. *American Journal of Psychology*, 1934, 46, 287-298.
- CRONBACH, L. J. Statistical methods for multi-score tests. *Journal of Clinical Psychology*, 1950, 6, 21-26.
- DANA, R. H. The validation of projective tests. *Journal of Projective Techniques*, 1962, 25, 182-186.
- DAWES, R. M. A note on base rates and psychometric efficiency. *Journal of Consulting Psychology*, 1962, 26, 422-424.
- DENNIS, W. The human figure drawings of bedouins. *Journal of Social Psychology*, 1960, 52, 209-219.
- DUDEX, F. J. Determining "chance success" when a specific number of items are sorted into discrete categories. *Journal of Consulting Psychology*, 1952, 16, 251-256.
- HAMLIN, R. M. The clinician as judge: Implications of a series of studies. *Journal of Consulting Psychology*, 1954, 18, 232-238.

- HAMMER, E. F. Critique of Swensen's "Empirical evaluations of human figure drawings." *Journal of Projective Techniques*, 1959, 23, 30-32.
- HUNT, W. A. The future of diagnostic testing in clinical psychology. *Journal of Clinical Psychology*, 1946, 2, 311-317.
- JONES, L. W., & THOMAS, C. B. Studies on figure drawings. *Psychiatric Quarterly Supplement*, 1961, 35, 212-261.
- KENDALL, M. G., & SMITH, B. B. Tables of random sampling numbers. In H. Arkin & R. R. Colton (Eds.), *Tables for statisticians*. New York: Barnes & Noble, 1950.
- LITTLE, K. B. Problems in the validation of projective techniques. *Journal of Projective Techniques*, 1959, 23, 287-290.
- LUBLIN, I., & LUBLIN, S. C. The Draw-A-Person Test: A Minimal Validity Investigation. Paper presented at the meeting of the Western Psychological Association, San Francisco, May 1967.
- MACHOVER, K. *Personality projection in the drawing of the human figure*. Springfield, Ill.: Charles C Thomas, 1949.
- MEEHL, P. E. *Clinical versus statistical prediction*. Minneapolis: University of Minnesota Press, 1954.
- MOGAR, R. E. Anxiety indices in human figure drawings: A replication and extension. *Journal of Consulting Psychology*, 1962, 26, 108.
- MOSTELLER, F., & BUSH, R. R. Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology*. Vol. 1. Cambridge, Mass.: Addison-Wesley, 1954.
- SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
- SCHNEIDER, L. I. Rorschach validation: Some methodological aspects. *Psychological Bulletin*, 1950, 47, 493-508.
- SECORD, P. F. A note on the problem of homogeneity-heterogeneity on the use of the matching method in personality studies. *Psychological Bulletin*, 1952, 49, 41-42.
- SHNEIDMAN, E. S. Suggestions for the delineation of validity studies. *Journal of Projective Techniques*, 1959, 23, 259-262.
- SUNDBERG, N. D. The practice of psychological testing in clinical services in the United States. *American Psychologist*, 1961, 16, 79-83.
- SWENSEN, C. H. Empirical evaluations of human figure drawings. *Psychological Bulletin*, 1957, 54, 431-466.
- VERNON, P. E. The matching method applied to investigations of personality. *Psychological Bulletin*, 1936, 33, 149-177.
- WALLON, E. J. *A study of criteria used to differentiate the human figure drawings of normals, neurotics, and psychotics*. (Doctoral dissertation, Purdue University) Ann Arbor, Mich.: University Microfilms, 1959. No. 59-418.

(Received February 8, 1968)

Editor's Note

The Editor has invited the following discussion by Emanuel Hammer.

DAP: BACK AGAINST THE WALL?

EMANUEL F. HAMMER¹

New York, New York

Methodological and conceptual criticism of Wanderer's study, "Validity of Clinical Judgments Based on Human Figure Drawings," revolves around (a) correlation procedures, (b) sampling errors, (c) arbitrary categorization of data, (d) overlap between diagnostic groups employed, (e) the insufficiency of fractionated material as projective samples, and, most importantly (f) clinical considerations.

Has Wanderer's study cracked open the hard nut of the question of the basic validity of the Draw-A-Person (DAP) technique?

As a research question runs the gauntlet through evidence and counterevidence, one tries to make sense of the resulting confusion. Wanderer's study joins those on the side which have failed to demonstrate that psychologists can offer valid diagnoses on the basis of figure drawings (Sherman, 1958a, 1958b; Sipprelle & Swensen, 1956; Swensen, 1957; Whitmyre, 1953). Along with this, several of Machover's (1949) interpretive hypotheses have emerged bearing evidence suggesting their lack of validity (Blum, 1954; Fisher & Fisher, 1950; Grams & Rinder, 1958; Hammer, 1954; Reznikoff & Nicholas, 1958; Ribler, 1957).

Yielding a mixture of results, Hammer and Piotrowski (1953) found (a) a reassuring degree of both reliability and validity in the interpretation of aggression reflected in the House-Tree-Person drawings, and also, (b) that what variation there was among the clinicians' interpretations appears to be secondarily influenced by their own projections and/or degree of sensitivity.²

On the other hand, and in contradistinction to the findings of Wanderer, a number of studies have demonstrated clear-cut differences between the DAPs of "normal" and pathological groups (Anastasi & Foley, 1941; Berrien, 1935; Eigenbode, 1951; Goldworth, 1950; Gunzburg, 1954; Holzberg & Wexler, 1950; Hozier, 1959; Plaut & Crannell, 1955; Reznikoff & Tomblen, 1956; Schmidl-Waehner, 1942; Singer, 1950; Springer,

1941; Wexler & Holzberg, 1952). Steinmann (1952) devised a scoring system which was found by Graham (1955) to correlate .70 with the degree of pathology in psychosis. Hiler and Nesvig (1965) also demonstrated that psychiatric patients could be differentiated from normals on the basis of projective drawings. They validated the characteristics of pathology as reflected in the following qualities: "bizarreness," "distorted aspects," "incompleteness" and "transparency" of drawings. The most valid of the criteria was "bizarreness," which subsumed a drawing looking either "grotesque," "inhuman," "sinister," "sick," "ghoulish," "weird," or "gnomelike." In a cross-validation sample, it was found that 46% of the patients and only 2% of the control group produced a drawing which, by independent judgment, was "bizarre."

Griffith and Peyman (1965) found that eye and ear emphasis in the DAP was, to a statistically significant degree, associated with ideas of reference. Albee and Hamlin (1950) devised a scale which, while unable to differentiate between schizophrenics and anxiety neurotics, did reliably differentiate between these two groups and a normal one.

Wanderer's study is now added to the negative side of the mixed results which have proliferated over the years. As we evaluate his study, there are, however, several ways in which it appears the hypotheses were handicapped.

No significant correlation is found between the "expertness" of the judges and the validity of their sortings. With all the experts possessing national reputation, they quite evidently fall into a tight cluster with little of the spread necessary to allow a statistically meaningful correlation.

Totally disqualifying one facet of the study, the judges did not even see the same set of five cases. There is therefore neither common ground nor any defensible basis for appraising one expert against another in correlating validity and "degree of expertness." Surprisingly, scientific method was abandoned to a situation where one judge

¹ Appreciative thanks are warmly extended to John N. Buck and Robert L. Wolk for their fruitful comments and suggestions. Requests for reprints should be sent to Emanuel F. Hammer, 381 West End Avenue, New York, New York 10024.

² We might view the latter finding as raising the question of the advisability of a personal psychotherapeutic experience for the projective technique interpreter, as it has long been advocated by many for the psychotherapist.

may have had a relatively classical set of drawings while another a quite confusing set.

In Table 1, "Number of hits" is put into the categories of "0," "1," or "2 or more." This kind of lumping of the data seems rather arbitrary. Combining all "hits" of 2, 3, 4, and 5 into this last category tends to dilute the data favoring the hypothesis. The judge who attained five out of five correct sortings does thereby not contribute any more weight to the hypothesis than the one who attained only two out of five correct matchings.

Also, the fact that the judges were asked to sort one DAP set into each of five categories did not allow a judge to elect to call two of the cases psychotic when they looked that way to him. Thus, in this instance when a judge was wrong on his first choice he was automatically wrong for a 40% rather than a 20% assessment, that is, the second case felt to also fall in the category already used up now had to be placed in a different one. Had both drawings been allowed to be classified schizophrenic, it might here have resulted in only a 20% error. (Then, of course, the same handicap occurred, now on the next level, when, in going through the drawings a second time, the judges again had to put only one case in a category.)

Wanderer subtracts the influence of the conspicuous contribution to the validity of the sortings made by the mental defective cases. It is hard to understand why he did not merely rotate each category and use analysis of variance to solve this problem more neatly.

Wanderer quotes the present author, in reference to that rare but striking schizophrenic sign which may appear in a drawing, as saying that the most appropriate way to test its validity statistically would be to compare in what groups it occurs *when* it does occur. Wanderer then criticizes this as a procedure ignoring false negatives. His argument here would be very much like contending that hallucinations or delusions in a psychiatric patient cannot be used as suggestive of a psychotic process since they do not occur in *most* psychotics. Wanderer's misconception of "false negatives" would result in the most striking diagnostic clues being the very ones which "research" demonstrates should be ignored.³

To turn to Wanderer's sample, there is a question of overlap in the five groups (with a limited sample of 20 in each) which tends to contaminate the criteria. Experience in outpatient mental hygiene clinics is that there is a high proportion

of borderline and ambulatory schizophrenics in this population. Wanderer doesn't indicate how, or if, he sifted them out of his "neurotic" (mental hygiene clinic) sample. Additionally, we know that intense homosexual conflicts are frequently part of the psychotic, as well as the neurotic, picture; and this would clearly overlap with the DAPs of the homosexual group. In fact, the *over* homosexual group, having come (at least comparatively) to peace with needs in this direction, may flood their drawings with *less* of the homoerotic than would the neurotic and psychotic groups, patients in whom the problem areas press insistently forward in their projective responses.

When the comparison is between homosexual and normal groups, Machover, Puzzo, Machover, and Plumeau (1959) demonstrate statistical significance in the reliability and the validity of DAP differentiation "blindly" done: The average two-way total scale intercorrelation among three graduate-student judges who each independently scored the protocols was found to be .83. The efficacy of differentiating male homosexual from nonhomosexual controls was significant beyond the .001 level of confidence.

As to how many of Wanderer's "homosexual" group may have been neurotic or psychotic, we do not know. There was no procedure to assay or rule these out.

The complications of overlapping of groupings may be seen more clearly in the examples that (A) a neurotic depression and (B) a psychotic depression, may look more alike than different in a projective sampling, particularly of just two drawings obtained. The same may be true of (C) a neurotic paranoid integration and (D) a paranoid type of schizophrenic. In fact, on this brief sample, A, looking more like B than like its official category mate C, and C looking more like D than like its experimentally defined companion A would mean that the depressive and paranoid dimensions, respectively, would emerge but not be differentially related to whether their presence was of neurotic or psychotic proportions. When such is the case, we could scarcely move toward the conclusion that the drawings here provided nothing diagnostically valid. The same point may be made when the major drawing theme is the pivotal one of strong dependency needs—whether, in line with Wanderer's groups, on the part of (a) a "normal" subject sustained by a marriage to a warm, maternal woman, (b) a "neurotic" heavy drinker, (c) a "homosexual" whose preferred erotic activity is performing fellatio, or (d) an institutionalized patient responding to the dependency support provided by the hospital. As Hutt (1968) points out, "Projective tests seem

³ "Some circumstantial evidence" says Thoreau, "is very strong, as when you find a trout in the milk."

to be especially valuable in assessing covert factors. . . . They are less successful in delineating or predicting overt behavior which is likely to be the end-product of many intervening factors."

Wanderer's study rests on a basic assumption of mutually exclusive pigeon holes for people. His data, however, if viewed from the broad perspective, stand as a banner to the refusal of complex human beings to fit into neat categories (at least on brief segments of behavior).

As far back as 1947, Tomkins suggested that the psychiatric entities "do not, in our opinion, represent homogenous entities even at the level of symptomatology." Schafer (1954), too, speaks of

researchers who naively expect that dumping all patients described as "paranoid" (or "anxious," "schizophrenic," "well-adjusted," etc.) into one group will consistently yield highly instructive (results) . . . and whose conception of test theory and research stops right there [p. 285].

The fact that the schizophrenic group, in Wanderer's study, was misdiagnosed more often than all of the other groups is consistent with the view that there is not a schizophrenic group per se. A proliferating of more varied types is reflected in Beck's (1954) expansion of four to six subtypes of schizophrenia as reflected in the Rorschach. Evelyn Hooker makes the point more broadly in asserting that there is no such entity as *the* homosexual. In describing some of the "worlds" of homosexuals seen in one large city, she reports no common pattern of personality structure, much less dynamics, in homosexuals. Willis (1967) similarly found no similarity among homosexuals other than in their choice of sex object.

The mentally defective group was undoubtedly diagnosed validly because of its clear-cut lowered intellectual and conceptual qualities. But it may also have been easier to diagnose because mental defectives are a more distinct entity whose members more intimately share a common denominator with less spread and variability. And where each of the other four groups had crossover features with each other, this one group had few or none (at least in one direction, that is, probably not a single "normal," "neurotic," "psychotic," or "homosexual" was also mentally defective).

Shaffer and Schafer are, I think, both right; Lawrence Shaffer (quoted by Wanderer in the paper under discussion) partly so, and Roy Schafer (1954) more so. We do, to be sure, tend to read *into* projective material from what we know of the case. But with this providing the frame of reference, we then can read *from* the

drawings significantly further. If—as in actual practice (operating neither "blindly" nor employing a mere splinter of the battery)—we know the patient to be in a psychotic depression and we find no affect, depressive or otherwise, in the projective drawings (or elsewhere in the projectives), we may extend from the known into the previously not known. We can then understand that a state of emotional dehydration may be present, and may be here employed as a last-ditch defense against the intolerable despondency and despair. The blandness may involve defensive denial of the inner experience. To follow the thread into the tissue of the understructure, we can then link up surface and subsurface so that it makes sense of the apparent inconsistency. If an individual offers a drawing production seeped in sadism and gore (and his Rorschach and TAT is similarly so) but behaviorally he is mild and meek, do we discard the projective level data as invalid, or do we deduce that this man's surface behavior is maintained at the price of considerable repression, inner strain, and tension?

In these cases, we can see illustrated the principle that a small projective sample may be insufficient to convey the central diagnostic quality of the patient, but may nevertheless still contribute deeper understanding to the personality picture.

Roy Schafer (1954) vigorously underscores the dangers of item interpretation and emphasizes a reliance instead on themes which extend across many projective products. Similarly, in the present writer's *The Clinical Application of Projective Drawings* (Hammer, 1958) may be found, prominently asserted:

In actual clinical practice, the dangers of basing interpretative deductions on isolated bits of data are obvious. In practice, confirmation of interpretative speculations on the basis of one drawing must be checked against not only the other drawings, but the entire projective battery, the case history, the clinical impression gleaned during the interview with the subject and all other available information [p. 49].

An experimental investigation which takes a mere chip off this procedure bears little resemblance to clinical practice and, as a study, has built into it only the slimmest chance of surviving its handicaps.

Another basic consideration has always been S's perceptions of the purpose of the test and his particular reason for the examination, a highly determining variable coloring S's style and responses (Cronbach, 1946; Hutt, Gibby, Milton, & Potthurst, 1950; Murstein, 1965; Tolor, 1968)—crucial material unavailable in "blind" assessment

studies. In accord with this, Tolor (1968) speaks of the "situational factors (which) can exert great influence on all types of performance, and must be taken into account." Tolor then goes on to join Buck (1966) and Machover (1949) in underscoring the importance of the patient's verbal associations to his own drawings, data whose absence in Wanderer's study fractionates the already fractionated material further. As Tolor (1968) advocates,

Place greater weight on associations produced in response to the drawings and do not depend solely on (the graphic projection). . . . Enlist the patient's assistance in establishing interpretive hypotheses. The mode used by a patient to reconcile disparate performance, to rationalize behavior, and to gain insight represents important data for interpretation [p. 226].

A projective device, by its nature, has a buckshot quality. Because of its uncertain aim, a projective technique cannot be appraised on the basis of whether or not it invariably focuses upon the same dimension in each *S*. Sometimes in the DAP a diagnostic hint emerges, but sometimes a personality need, sometimes a psychodynamic clue, and sometimes an expressive style of personality. Two drawings of a person do not constitute a system in itself, but only a small facet of an unfolding system of responses maintaining itself in contact with other systems of the patterning.⁴

Only a minority of the experts employed as judges in Wanderer's study regard the DAP as a "primary tool." No one uses the DAP alone. It was never intended by Machover, or anyone else, as anything more than a supplement, a graphic adjunct to the verbal techniques. Therefore, Wanderer's statement,

but one could still wonder about how much confidence could be placed in an instrument which, *used alone*, would send a large proportion of schizophrenics merrily on their way, since they would be identified as anything but schizophrenics [p. 149, *italics supplied*],

reflects an astonishing failure to understand the role and utility of projective drawings in the battery.

Overall, Wanderer's study demonstrates what we know clinically: that to order and integrate findings from the various sources of data in the testing situation, the clinician cannot rely solely

⁴ Sometimes, for example, with one schizophrenic the pathology comes through on one technique and with another on a different technique, and frequently only in the relationship among the techniques. This is, of course, the reason we employ multiple measurements from diverse modalities.

on any small, compartmentalized sampling. A 5- or 10-minute projective yield cannot, alone and when "blindly" interpreted, serve as the basis for a diagnosis.⁵ Interpreting a personality must be in terms of its inner unity and consistency across some degree of spread; only thereby may we, at times, reveal the man.

Anything less belongs with parlor tricks. It is in line with this that we may view the especially pertinent study of Caligor (1962). He found that paranoid trends could be detected in only 25% of a group of paranoid schizophrenics when one drawing of a human figure was employed. He returned to the problem, employing a set of eight such drawings, and found that 85% of the cases could be validly detected on the basis of the more extensive drawing projections.

Yet in spite of its defects—statistically, in terms of sampling, experimental design, and clinical sophistication—the impact of Wanderer's study still retains some voice which speaks to the intellect. As scientist-practitioners, we cannot ignore our everyday clinical evidence, nor can we, on the other hand, ignore the things that go on inside the limitations of the tight research design. In a field which is after truth rather than in much possession of it as yet, we must draw our working guidelines simultaneously from the microscopic analysis of bits and pieces of human behavior and also from our broader experience of its totality.

Clinicians may respond in one of three ways. They may ignore studies such as Wanderer's. But that makes little sense. They may, on absorbing the encounter with Wanderer's experiment, drop the DAP from the battery. But that would make little sense, too, unless they were using the DAP blindly, without the rest of the battery, and merely to establish a diagnostic category. They may, harmonizing the implications of Wanderer's with that of Caligor's study, resonate to the limitations of a two-drawing technique. They may then strengthen the projective drawing assessment with an expansion in width and depth.⁶ And that, to this clinician, makes more sense.

⁵ At least if the sample is of drawing products. Can an equally brief sample of Rorschach behavior or of TAT stories (say, the first two cards of the former, or the first story or two of the latter), or of MMPI answers, or of interview do so? We don't know, but few, if any of us, would expect it to do so.

⁶ The writer, for example, does not employ, or teach, the DAP or the H-T-P but rather an integrated drawing battery including the combined use of these two techniques, the H-T-P-P, both achromatic and chromatic (in pencil and then again in crayon), its searching Post-Drawing Interrogation

REFERENCES

- ALBEE, G. W., & HAMLIN, R. M. Judgment of adjustment from drawings: The applicability of rating scale methods. *Journal of Clinical Psychology*, 1950, 6, 363-365.
- ANASTASI, A., & FOLEY, J. P., JR. A survey of literature on artistic behavior in the abnormal: Experimental investigations. *Journal of Genetic Psychology*, 1941, 23, 187-237.
- BECK, S. J. *The six schizophrenias*. New York: American Orthopsychiatric Association, 1954.
- BERRIEN, F. K. A study of the drawings of abnormal children. *Journal of Educational Psychology*, 1935, 26, 143-150.
- BLUM, R. H. The validity of the Machover DAP technique. *Journal of Clinical Psychology*, 1954, 10, 120-125.
- BUCK, J. N. *The House-Tree-Person Technique, revised manual*. Beverly Hills, Calif.: Western Psychological Services, 1966.
- CALIGOR, L. The detection of paranoid trends by the 8 Card Re-Drawing Test. *Journal of Clinical Psychology*, 1952, 8, 397-401.
- CRONBACH, L. J. Response sets and test validity. *Educational and Psychological Measurement*, 1946, 6, 475-494.
- EIGENBODE, C. R. Effectiveness of the Machover signs and others in differentiating between a normal group and a schizophrenic group by use of the projective drawing test. Unpublished master's thesis, George Washington University, 1951.
- FISHER, S., & FISHER, R. Test of certain assumptions regarding figure drawing analysis. *Journal of Abnormal and Social Psychology*, 1950, 45, 727-732.
- GOLDWORTH, S. A. A comparative study of the drawings of a man and a woman done by normal, neurotic, schizophrenic and brain damaged individuals. Unpublished doctoral dissertation, University of Pittsburgh, 1950.
- GRAHAM, S. Relation between histamine tolerance, visual autokinesis, Rorschach human movement, and figure drawing. *Journal of Clinical Psychology*, 1955, 11, 370-373.
- GRAMS, A., & RINDER, L. Signs of homosexuality in human figure drawings. *Journal of Consulting Psychology*, 1958, 22, 394.
- GRIFFITH, A., & PEYMAN, D. Eye-ear emphasis in the DAP Test as indicating ideas of reference. In B. I. Murstein, *Handbook of projective techniques*. New York: Basic Books, 1965.
- GUNZBURG, H. C. Scope and limitations of the Good-enough drawing test method in clinical work with mental defectives. *Journal of Clinical Psychology*, 1954, 10, 8-15.
- HAMMER, E. F. Relationship between diagnosis of psychosexual pathology and the sex of the first drawn person. *Journal of Clinical Psychology*, 1954, 10, 168-170.
- HAMMER, E. F. *The clinical application of projective drawings*. Springfield, Ill.: Charles C Thomas, 1958.
- HAMMER, E. F., & PIOTROWSKI, Z. A. Hostility as a factor in the clinician's personality as it affects his interpretation of projective drawings. *Journal of Projective Techniques*, 1953, 17, 210-216.
- HILER, E., & NESVIG, D. Evaluation of criteria used by clinicians to infer pathology from figure drawings. *Journal of Consulting Psychology*, 1965, 29, 520-529.
- HOLZBERG, J. D., & WEXLER, M. The validity of human form drawings as a measure of personality deviation. *Journal of Projective Techniques*, 1950, 14, 343-361.
- HOZIER, A. On the breakdown of the sense of reality: A study of spatial perception in schizophrenia. *Journal of Consulting Psychology*, 1959, 23, 185-194.
- HUTT, M. L. Psychopathology, assessment and psychotherapy. In A. I. Rabin (Ed.), *Projective techniques in personality assessment*. New York: Springer, 1968.
- HUTT, M. L., GIBBY, R. G., MILTON, E., & POTT-HURST, K. The effect of varied experimental "sets" on Rorschach test performance. *Journal of Projective Techniques*, 1950, 14, 181-187.
- MACHOVER, K. *Personality projection in the drawing of the human figure*. Springfield, Ill.: Charles C Thomas, 1949.
- MACHOVER, S., PUZZO, F., MACHOVER, K., & PLUMEAU, F. An objective study of homosexuality in alcoholism. *Quarterly Journal of Studies in Alcoholism*, 1959, 20, 528-542.
- MURSTEIN, B. I. *Handbook of projective techniques*. New York: Basic Books, 1965.
- PLAUT, E., & CRANNELL, C. W. The ability of clinical psychologists to discriminate between drawings by deteriorated schizophrenics and drawings by normal subjects. *Psychological Reports*, 1955, 1, 153-158.
- REZNIKOFF, M., & NICHOLAS, A. L. An evaluation of human-figure drawing indicators of paranoid pathology. *Journal of Consulting Psychology*, 1958, 22, 395-397.
- REZNIKOFF, M., & TOMBLEN, D. The use of human figure drawings in the diagnosis of organic pathology. *Journal of Consulting Psychology*, 1956, 20, 467-470.
- RIBLER, R. I. Diagnostic prediction from emphasis on the eye and the ear in human figure drawings. *Journal of Consulting Psychology*, 1957, 21, 223-225.
- SCHAFER, R. *Psychoanalytic interpretation of Rorschach testing*. New York: Grune & Stratton, 1954.
- SCHMIDL-WAEHNER, R. Formal criteria for the analysis of children's drawings. *American Journal of Orthopsychiatry*, 1942, 17, 95-104.
- (Buck, 1966), the Draw-A-Person-In-The-Rain modification which attempts to elicit clues to the self-concept under conditions symbolizing environment stress, the Draw-An-Animal approach useful for disclosing the biological side of the biosocial coin, the Eight-Card-Redrawing Test which digs down into the deeper layers of S's psychosexual identification, the Unpleasant Concept Test, and with children, the Draw-A-Family procedure.

- SHERMAN, L. J. The influence of artistic quality on judgments of patient and non-patient status from human figure drawings. *Journal of Projective Techniques*, 1958, 22, 338-340. (a)
- SHERMAN, L. J. Sexual differentiation of artistic ability? *Journal of Clinical Psychology*, 1958, 14, 170-171. (b)
- SINGER, R. H. A study of drawings produced by a group of college students and a group of hospitalized schizophrenics. Unpublished master's thesis, Pennsylvania State University, 1950.
- SIPRELLE, C. N., & SWENSEN, C. H. Relationships of sexual adjustment to certain sexual characteristics of human figure drawings. *Journal of Consulting Psychology*, 1956, 20, 197-198.
- SPRINGER, N. N. A study of the drawings of maladjusted and adjusted children. *Journal of Genetic Psychology*, 1941, 58, 131-138.
- STEINMANN, K. The validity of projective technique in the determination of relative intensity in psychosis. Unpublished doctoral dissertation, School of Education, New York University, 1952.
- SWENSEN, C. H. Empirical evaluation of human figure drawings. *Psychological Bulletin*, 1957, 54, 431-466.
- TOLOR, A. The graphomotor techniques. *Journal of Projective Techniques and Personality Assessment*, 1968, 32, 222-228.
- TOMKINS, S. S. *The Thematic Apperception Test*. New York: Grune & Stratton, 1947.
- WEXLER, M., & HOLZBERG, J. D. A further study of the validity of human form drawings in personality evaluation. *Journal of Projective Techniques*, 1952, 16, 249-251.
- WHITMYRE, J. W. The significance of artistic excellence in the judgment of adjustment inferred from human figure drawings. *Journal of Consulting Psychology*, 1953, 17, 421-424.
- WILLIS, S. *Understanding and counseling the male homosexual*. New York: Little, Brown, 1967.

(Received July 26, 1968)

CREATIVITY, AWARENESS, AND LIKING

RUSSELL EISENMAN¹

Temple University

As a follow-up to a previous investigation of Barron's creativity theory, and in order to extend previous findings, 302 college students took an unusual uses test, the Barron-Welsh Art Scale (BWAS), and a polygon measure of complexity-simplicity. In addition, they rated how much they liked being in the study on a 1-7 scale, and stated what they thought was the purpose of the research. The fluency and originality scores, derived from the unusual uses test, showed very high intercorrelations (p 's $< .001$), while the BWAS and polygon measures were intercorrelated with each other ($r = .55$, $p < .001$) and with originality and fluency. Although the results seem partially due to method variance, the correlations suggest, in contrast to several previous findings, that there are creativity measures which show a reasonable degree of interrelationship, thereby justifying subsuming them under the construct "creativity." The importance of awareness and liking are also considered.

A recent study (Eisenman, 1968a) found support for Barron's (1953, 1963) view that complexity-simplicity is an important personality dimension. However, Eisenman (1968a) found only modest support for the notion that complexity-simplicity can be conceived of as a creativity variable. Since Barron (1963) has presented several studies relating preference for complexity to creativity, the failure of the Eisenman study to show convincing support for the complexity-creativity linkage raises some problems. First, was Eisenman really measuring creativity with his polygon measure of complexity-simplicity preference? Second, is there any relationship between Eisenman's polygon measure of complexity-simplicity and Barron's line drawings, which are known as the Barron-Welsh Art Scale (BWAS)? Third, do creativity measures show any degree of interrelationship with one another? This third point may be the crux of the matter, since failure to obtain significant intercorrelations suggests that one's various measures cannot be subsumed under the general rubric "creativity."

Numerous researchers (Barron, 1955; Eisenman & Robinson, 1967; Getzels & Csikszentmihalyi, 1966; Piers, 1968) have reported insignificant or small correlations among their various creativity measures. This lack of substantial correlation among creativity tests has

been seized upon by Wallach and Kogan (1965) who have criticized the well-known Getzels and Jackson (1962) creativity-IQ investigation because Getzels and Jackson's creativity measures correlated as highly with IQ tests as they did with one another. Other studies, in addition to the ones mentioned above, suffer from the lack of interrelationship among the creativity measures utilized. One focus of the present study was to see if tasks widely employed for measuring creativity showed any substantial degree of intercorrelation. To this end, a large number of Ss were tested with the BWAS, Eisenman's polygons for complexity-simplicity preferences, and expressed uses for three common objects. The uses test has typically been scored for originality (infrequency of response in the sample tested) and for fluency (total number of valid responses). Evidence (Christensen, Guilford, & Wilson, 1957; Wallach & Kogan, 1965) suggests that common responses tend to occur early in the response sequence, and the more original responses occur later. Thus, there is a sort of built-in correlation between fluency and originality; since the person who gives many responses is likely to exhaust his common responses and move into the more original ones. To correct for this, a fifth creativity score was employed which simply consisted of dividing number of original responses by total number of valid responses. The originality/fluency measure gives us an "originality average" in much the same way that a batting

¹ Requests for reprints should be sent to Russell Eisenman, Department of Psychology, Temple University, Philadelphia, Pennsylvania.

average is obtained in baseball via the formula hits/times at bat.

Psychological researchers are becoming increasingly aware of Ss' feelings about the research in which he participates, both because Ss' expectations may influence the results and because of ethical concerns. Two such areas of "subject's reaction" were considered in the present study: awareness and liking. Awareness has been frequently investigated in verbal conditioning studies, but greatly ignored in creativity research. It seemed reasonable to suppose that S's awareness may be correlated with some performance, either because the aware S can better give E the results that E "wants," or because awareness might reflect general brightness or interest. In summary, the present study was designed as a follow-up to Eisenman's (1968a, 1968b) findings, and to consider aspects of creativity research which have been neglected (awareness and liking) or which have often failed to receive substantiation (interrelationships among creativity measures).

METHOD

Subjects

The Ss were 302 students from undergraduate classes at Temple University.

Procedure

The Ss were asked to participate in a psychological research study. No explanation was given for the instruments used, other than telling Ss that E was interested in preferences. The testing was done in large groups, with care being taken to see that no "cheating" occurred. All questions were answered as

briefly as possible, with a minimum of information. Most queries were handled by telling Ss, "it's up to you." Typically, the unusual uses test was administered first, with Ss told to list all the uses they could think of for a brick, a paper clip, and a pencil. Approximately half the Ss took the BWAS second and the polygons third, while the other Ss received the polygon preference measure second and the BWAS third. No order effects were found. The polygons were on photographs, with three each of 4-, 12-, and 24-point random shapes, and one each of 4-, 8-, and 10-point symmetrical shapes. A complexity score was obtained by subtracting total number of points on three least preferred shapes from total number of points on three most preferred shapes. Fluency scores were derived from the unusual uses test by obtaining the total number of valid responses given by S, while originality scores were obtained by crediting S with 1 point for originality every time his response occurred with a frequency of 1% or less among all responses in the present study. Scores across the three unusual uses items were combined for purposes of computing Pearson r 's. An assistant and E scored the unusual uses items for originality and fluency, and agreed on over 90% of the classifications. Since BWAS and the polygon measure are objectively scored, there was no need for two scorers.

At the conclusion of the unusual uses, BWAS, and polygon measures, Ss rated how much they liked being in the study, using a scale of 1-7 with 1 indicating the greatest dislike, and 7 indicating greatest liking. The Ss also wrote what they considered to be the purposes of the study, and Ss who said such things as "creativity," "originality," or "test of thinking" were considered aware. For computing correlations, aware Ss were assigned a score of 1, and unaware Ss a score of 0. Two raters achieved 96% agreement.

RESULTS

Creativity

Table 1 shows the intercorrelation of all seven variables. Fluency and originality are

TABLE 1
INTERCORRELATION OF SEVEN VARIABLES

Variables	1	2	3	4	5	6	7
Fluency (1)							
Originality (2)	.85***						
Originality/Fluency (3)	.60***	.91***					
Complexity (4)	.19**	.36***	.37***				
Barron-Welsh Art Scale (5)	.27***	.38***	.39***	.55***			
Liked the study (6)	.45***	.45***	.40***	-.07	.02		
Aware of purpose of the study (7)	-.03	.13*	.26***	.05	.03	.45***	

Note.—For liking the study, Ss rated the study from 1 to 7, with 7 indicating the greatest liking. For awareness, Ss were classified as 0 (unaware) or 1 (aware). Decimals have been omitted. $N = 302$.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

highly related to each other ($r = .85$, $p < .001$). However, the correlations of the fluency and originality measures with BWAS and the complexity-simplicity polygon measure suggest that originality is more strongly associated with these alleged creativity measures than is fluency. Preference for complexity is associated with BWAS score ($r = .55$, $p < .001$). The major value of the "originality average" obtained by the formula originality/fluency seems to be the reduction in the relationship with fluency from .85 for originality and fluency to .60 for originality/fluency and fluency. In other words, Ss obtain a high correlation of originality and fluency in part because the more fluent S has a greater chance to be original. Even with the correction of dividing fluency scores by originality scores, there is still a healthy correlation of .60 between fluency and the "originality average," suggesting that original Ss are also fluent.

Awareness and Liking

If the person did well on the unusual uses test, he tended to like the study. However, his performance on BWAS and the polygons was not related to his rated liking of the study. Awareness showed less of a relationship to other variables than liking. The Ss who were original showed a slight tendency to be aware of the nature of the study in contrast to Ss with lower originality. Neither fluency nor the two preference measures (BWAS and polygons) showed any relationship with awareness. Awareness was associated with liking the study ($r = .45$, $p < .001$).

DISCUSSION

In contrast to many previous studies, the present investigation yielded reasonable correlations among creativity measures, supporting the belief that these measures are dealing with a common concept of "creativity." However, the strongest correlations were found between fluency and originality measures from the unusual uses test, which can be explained in part as method variance (Campbell & Fiske, 1959) since a single response leads to more than one score. The r of .55 between BWAS and complexity-simplicity preferences indicates that previous studies (Eisenman, 1968b) which used polygon preferences to

test extensions of Barron's (1963) theory were accurate insofar as complexity₁ has some overlap with complexity₂. Complexity₁ refers to preference for complex line drawings in what developed into the BWAS, while complexity₂ is preference for polygon complexity. Of course, these two measures are far from being identical. The r of .55 accounts for less than one-third of the variance, and previous research (Eisenman & Rappaport, 1967) has suggested at least one way in which the BWAS and the polygon preference measures are different. Nevertheless, there is some overlap which is a welcomed change in creativity research, since many previous studies have found less of an interrelationship.

The patterning of the correlations with the BWAS and with complexity-simplicity preferences is highly similar. Correlations with BWAS are never greatly different than the corresponding correlations with complexity-simplicity, which further strengthens the belief that these two alleged creativity measures have much in common.

It is probably a testimony to the subtlety of the preference measures that they showed no significant correlations with either awareness or liking. On the other hand, some Ss seem capable of correctly hypothesizing that the unusual uses test is aimed at measuring creativity or some kind of cognitive function. Perhaps the ability to figure out what the researcher is investigating reflects general brightness or interest.

The answer would seem to be "yes" for all three questions posed in the introduction. The present study suggests that (a) polygon complexity-simplicity has claim as a creativity measure; (b) complexity-simplicity and BWAS are related; (c) there are creativity measures which show some reasonable degree of intercorrelation.

REFERENCES

- BARRON, F. Complexity-simplicity as a personality dimension. *Journal of Abnormal and Social Psychology*, 1953, 21, 287-297.
- BARRON, F. The disposition toward originality. *Journal of Abnormal and Social Psychology*, 1955, 51, 478-485.
- BARRON, F. *Creativity and psychological health*. Princeton, N. J.: Van Nostrand, 1963.
- CAMPBELL, D. T., & FISKE, D. W. Convergent and discriminant validation by the multitrait-multi-

- method matrix. *Psychological Bulletin*, 1959, 56, 81-105.
- CHRISTENSEN, P. R., GUILFORD, J. P., & WILSON, R. C. Relations of creative responses to working time and instruction. *Journal of Experimental Psychology*, 1957, 53, 82-88.
- EISENMAN, R. Personality and demography in complexity-simplicity. *Journal of Consulting and Clinical Psychology*, 1968, 32, 140-143. (a)
- EISENMAN, R. The psychology of modern art. In G. P. Powers & W. Baskin (Eds.), *New outlooks in psychology*. New York: Philosophical Library, 1968. (b)
- EISENMAN, R., & RAPPAPORT, J. Complexity preference and semantic differential ratings of complexity-simplicity and symmetry-asymmetry. *Psychonomic Science*, 1967, 7, 147-148.
- EISENMAN, R., & ROBINSON, N. Complexity-simplicity, creativity, intelligence, and other correlates. *Journal of Psychology*, 1967, 67, 331-334.
- GETZELS, J. W., & CSIKSZENTMIHALYI, M. The study of creativity in future artists: The criterion problem. In O. J. Harvey (Ed.), *Experience, structure and adaptability*. New York: Springer, 1966.
- GETZELS, J. W., & JACKSON, P. W. *Creativity and intelligence*. New York: Wiley, 1962.
- PIERS, E. V. Adolescent creativity. In J. F. Adams (Ed.), *Understanding adolescence*. Boston: Allyn & Bacon, 1968.
- WALLACE, M. A., & KOGAN, N. *Modes of thinking in young children*. New York: Holt, Rinehart & Winston, 1965.

(Received March 11, 1968)

REPRESSION-SENSITIZATION: ITS RELATION TO ADJUSTMENT AND SEEKING PSYCHOTHERAPY AMONG COLLEGE STUDENTS¹

MARK H. THELEN²

University of Missouri at Columbia

This study investigated the variables of adjustment and voluntarily approaching psychotherapy as each relates to repression-sensitization (*R-S*). A group of college males and females who sought psychotherapy (*T*) were compared on the *R-S* scale with a similar group which had not sought psychotherapy (*NT*). The results indicate that the *T* group scored significantly higher on the *R-S* scale than the *NT* group. In addition, the *T* group was found to be more sensitizing than the *NT* group even when the two groups were equated on a paper-and-pencil measure of adjustment. Among males in the *T* group, those terminating therapy after fewer than four sessions were less sensitizing than those continuing in therapy.

A significant point of interest in research on repression-sensitization (*R-S*) concerns the relationship of this personality construct to adjustment. The *R-S* dimension is presumably indicative of an individual's tendency to approach or avoid stress (Byrne, 1964). Initially, it was expected that scoring at either extreme on the *R-S* dimension would be indicative of maladjustment (Byrne, 1961). This led to a flurry of studies relating the *R-S* scale (Byrne, 1961) and the revised *R-S* scale (Byrne, Barry, & Nelson, 1963) to paper-and-pencil measures of adjustment (Byrne, 1961; Byrne et al., 1963; Byrne, Golightly, & Sheffield, 1965; Joy, 1963; Tucky & Grigg, 1964). In general, however, these investigations suggest that repressors are better adjusted than sensitizers. However, these studies have two serious limitations which could account for the findings. First, many of the studies employed measures of adjustment which contained items also contained in the *R-S* scale. Byrne et al. (1965) point this out; however, their study is subject to the same limitation because of item overlap between the *R-S* scale and the California Personality Inventory (CPI), which they used as a mea-

sure of adjustment. Second, repressors are likely to deny that they have problems and thus appear well adjusted on a paper-and-pencil measure.

In order to correct for the above limitations in assessing adjustment, Feder (1967) studied hospitalized psychiatric patients, and Tempone and Lamb (1967) studied patients at a mental health center. Both studies found that sensitizers tended to be in the psychiatric group and repressors in the control-adjusted group. The present study attempts to ascertain if a similar pattern exists in a population of college students seeking psychotherapy at a university outpatient mental hygiene clinic.

It is likely, however, that psychiatric outpatients are not only more maladjusted than controls but also respond to their problems differently. Therefore, one person may seek psychotherapy, while another, equally maladjusted person, does not. People seeking therapy are, in effect, approaching their problems and should score toward the sensitization end of the *R-S* dimension. The present study investigates the *R-S* dimension in such a population while controlling for adjustment.

Related to the issue of *R-S* and adjustment is the question of what the *R-S* scale really measures. A correlation of .87 between the *R-S* scale (a high score indicates sensitization) and the Taylor Manifest Anxiety Scale led Golin, Herron, Takota, and Reineck (1967) to suggest that the two tests may be interchangeable. Lower but clearly significant correla-

¹ The author is grateful to June Chance, Robert Dolliver, Douglas Ewing, Joseph Kuncze, Charles Krauskopf, Al Landfield, and Noel Maze for their constructive comments and assistance during various phases of this study.

² Requests for reprints should be sent to Mark H. Thelen, Department of Psychology, University of Missouri, Columbia, Missouri 65201.

tions are reported between the *R-S* scale, and Social Desirability and Acquiescence Response Set (Feder, 1967). Also, Lefcourt (1966) reported that *Ss* scoring low on the *R-S* scale interpreted the scale as a measure of mental illness, whereas sensitizers perceived the scale as an indicator of honesty with one's self. Thus, Lefcourt writes,

the question may be raised as to whether the *R-S* scale research is at all involved in the same area of interest as studies dealing with behavior referents of repression, isolation, or sensitization [1966, p. 448].

It appears that the behavioral referents of the *R-S* scale are under question despite the historical roots of the scale (Bruner & Postman, 1947). This study examines the behavior of approaching personal problems by seeking therapy in relation to the *R-S* scale. Since the *R-S* scale appears to be related to adjustment, this study attempts to equate Therapy (T) and Nontherapy (NT) *Ss* for adjustment and then determine if T *Ss* score higher than NT *Ss* on the *R-S* scale. The present study also explores the relationship between *R-S* and early termination of psychotherapy.

With a college population, including students who voluntarily sought psychotherapy, the following predictions are examined:

1. *Ss* in the T group are expected to score higher on the *R-S* scale (sensitization) than *Ss* in the NT group.

2. With adjustment controlled by matching *Ss*, the T group is still expected to score higher on the *R-S* scale than the NT group.

3. T *Ss* who terminate before the fourth session will score lower on the *R-S* scale than T *Ss* who continue in therapy for a larger number of meetings.

Because of their significance in this study and in previous research (Byrne et al., 1965; Joy, 1963), the correlations between the *R-S* scale and the CPI scales will also be reported.

METHOD

All the college undergraduate students who voluntarily sought psychotherapy at a university mental hygiene clinic during approximately a 3-month period were asked to participate in the study, providing they had not received psychotherapy during the previous 2 years. Of those asked to participate, approximately 75% completed the testing giving a T group of 63 *Ss* (27 males and 36 females).

The *Ss* for the NT group ($N = 112$; 61 males, 51 females) were obtained from introductory psychology and personal adjustment classes, excluding those who had received psychotherapy during the 2 previous years. Eighty-two percent of the NT group and 73% of the T group were either freshmen or sophomores. The remaining students in each group were either juniors or seniors.

The CPI and the Minnesota Multiphasic Personality Inventory (MMPI) were completed by *Ss* under the usual conditions of written instructions. All *Ss* in the T group completed the testing before their second therapy session. The CPI was employed as a measure of adjustment and the *R-S* scale (Byrne et al., 1963) was obtained from the MMPI. Although there is some item overlap between the *R-S* scale and nearly all of the CPI scales, this overlap would tend to make the T and NT groups similar on the *R-S* scale, not different as predicted.

For the analysis having to do with the number of therapy sessions attended, the figures excluded the intake session which is a brief diagnostic interview and often results in assignment to another therapist. The criterion for early termination was attendance at no more than three therapy sessions. It was reasoned that a satisfactory resolution of the clients' problems would not likely occur in fewer than four therapy sessions.

RESULTS

Prior to presenting the results pertaining to the hypotheses, the relationship between the *R-S* and the CPI scales is reported in Table 1. As may be seen, the *R-S* scale is significantly and negatively correlated with 15 of the 18 CPI scales for both the T and NT groups. However, certain problems, discussed in the next section, render these findings difficult to interpret.

The first hypothesis called for a comparison of the T and NT groups on the *R-S* scale. The T group obtained a mean *R-S* score of 62.03, which, when compared with the mean *R-S* score of 42.43 for the NT group, gives a t of 5.49 ($p < .001$, one-tailed). Thus the T group is significantly more sensitizing than the NT group. No sex differences in *R-S* scores were observed in the T group ($t = .91$) or in the NT group ($t = .32$).

The second prediction was that the T group would score higher on the *R-S* scale with adjustment controlled. To equate for adjustment, as many *Ss* as possible in the T group were matched with an *S* in the NT group on the basis of an identical CPI scale score. This was done separately for each of the 18 CPI

TABLE 1

CORRELATIONS BETWEEN THE R-S AND CPI SCALES FOR THE T AND NT GROUPS

CPI scales	T group	NT group
Sociability	-.48**	-.45**
Femininity	-.14	+.05
Dominance	-.33**	-.27**
Achievement via Independence	-.48**	-.39**
Capacity for Status	-.54**	-.51**
Communality	-.26*	-.34**
Self-Control	-.67**	-.61**
Psychological-Mindedness	-.45**	-.39**
Self-Acceptance	-.15	-.12
Responsibility	-.28*	-.39**
Sense of Well-Being	-.76**	-.78**
Intellectual Efficiency	-.76**	-.64**
Achievement via Conformance	-.65**	-.68**
Tolerance	-.70**	-.67**
Good Impression	-.66**	-.66**
Flexibility	-.12	-.04
Social Presence	-.55**	-.40**
Socialization	-.38**	-.41**

Note.— \rightarrow levels are two-tailed. For the T group, $N = 63$; for the NT group, $N = 112$.

* $p < .05$.
 ** $p < .01$.

scales.^a The results of the matched-paired comparisons, reported in Table 2, show significant differences between the T and NT groups on R-S for all 18 CPI scales and support the second hypothesis. Thus the T group, in comparison with the NT group, tends toward sensitization even when the groups are equated on a measure of adjustment. These strong differences occurred despite many instances of item overlap between the R-S scale and the CPI.

The third hypothesis concerns the relationship between R-S and early termination of psychotherapy. Because males and females may view early termination differently, comparisons were made for each sex. Table 3 shows that among males early terminators tend less toward sensitization than the remainder of the T group. No significant difference was found among the female Ss, although a similar trend is apparent.

^a Despite the large S pool in the NT group, it was not possible to find a match for every T S. Over the 18 CPI scales, the number of Ss with no match and therefore not included in the analysis, ranges from 3 to 19 with a mean of 9.6. In over half of these instances an appropriate match was found but the S had already been matched with another S in the T group.

DISCUSSION

The significant negative correlation between the R-S and most of the CPI scales is similar to findings reported by Byrne et al. (1965) and Joy (1963). The Ss scoring toward sensitization tend to score lower on many of the CPI scales. However, the particular interpretation of this apparently predictable observations is not so clear. As Byrne et al. (1965) pointed out, repressors might well appear better adjusted on a paper-and-pencil measure because they tend to deny that they have problems rather than because they are better adjusted. A second difficulty in interpreting these correlations, and one not mentioned by Byrne, is the item overlap between the R-S and most of the CPI scales. The extent of item overlap ranges as high as 9 of 31 items on the CPI Tolerance scale. Those scales having the greatest item overlap generally show high correlations with R-S (e.g., Sense

TABLE 2

COMPARISONS OF THE T GROUP WITH THE NT GROUP ON THE R-S SCALE WHEN Ss ARE MATCHED ON THE CPI SCALES

CPI Scale	Number of pairs	Mean R-S scores		<i>t</i>	<i>p</i> < *
		T group	NT group		
Sociability	44	58.9	46.6	2.63	.01
Femininity	55	62.8	40.9	5.90	.0005
Dominance	54	58.8	44.2	3.60	.0005
Achievement via Independence	55	60.3	42.7	4.34	.0005
Capacity for Status	60	61.9	46.6	4.14	.0005
Communality	60	62.5	43.6	4.67	.0005
Self-Control	48	58.2	40.1	4.82	.0005
Psychological-Mindedness	59	61.2	42.4	4.43	.0005
Self-Acceptance	59	61.0	42.9	4.20	.0005
Responsibility	56	59.1	45.6	3.31	.001
Sense of Well-Being	50	55.8	45.7	3.05	.0025
Intellectual Efficiency	49	56.8	45.2	3.50	.001
Achievement via Conformity	58	59.7	43.9	4.57	.0005
Tolerance	52	61.3	49.8	3.12	.0025
Good Impression	59	62.6	44.1	5.39	.0005
Flexibility	54	64.1	40.3	5.26	.0005
Social Presence	48	57.7	46.1	2.91	.005
Socialization	47	60.4	49.1	2.49	.01

* one-tailed.

TABLE 3

COMPARISON OF THERAPY Ss WHO ATTENDED THREE OR LESS SESSIONS WITH THOSE ATTENDING FOUR OR MORE SESSIONS ON THE R-S SCALE

Sex group	Number of sessions				<i>t</i>
	0-3		4 or more		
	<i>N</i>	<i>M</i>	<i>N</i>	<i>M</i>	
Males	8	55.12	19	69.42	1.72*
Females	11	54.73	25	61.84	1.17

* $p < .05$, one-tailed.

of Well-Being, 12 of 44; Social Presence, 13 of 56). Furthermore, most of the overlap items have opposite scorings on the *R-S* scale and the CPI; for example, those items scored for a true response on the *R-S* scale are nearly all scored for a false response on the CPI. This pattern of scoring would lead one to expect the negative correlations which were observed (see Table 1).

Consistent with the findings of Feder (1967) and Tempone and Lamb (1967), this study indicates that *R-S* is related to a broad non-paper-and-pencil measure of adjustment. Thus people seeking psychotherapy at a university outpatient mental health clinic tend toward sensitization in comparison with a comparable control group not seeking psychotherapy. In addition, however, the findings of the present study strongly suggest that the *R-S* scale is more than a measure of adjustment. Even when T and NT Ss were equated for adjustment on the CPI, Ss in the T group tended to be sensitizers and Ss in the NT group tended to be repressors.⁴ As evidenced by the behavior of seeking therapy, this study suggests that the *R-S* scale does have the behavioral referents, such as avoidance and sensitization, which have been questioned by Lefcourt (1966). Perhaps the *R-S* scale measure "adjustment" as well as the tendency to approach or avoid stress. Such a relationship does not make the terms interchangeable and certainly does not reduce the value of the

⁴ Since considerable item overlap exists between the 18 CPI scales, they may not be seen as independent measures.

R-S concept. When indicated, future research with the *R-S* scale might profitably control for adjustment to insure that observed differences are not a function of adjustment variables tapped by the *R-S* scale.

The present study observed no sex differences in *R-S* in either the T or NT groups. Most previous studies report similar findings or simply do not raise the question. This is reflected in a review article by Byrne (1964) in which no section is given to sex differences on the *R-S* dimension. However, the present study did obtain different results for males than females as regards the relationship between *R-S* and early termination of psychotherapy. Merbaum and Badia (1967) and Mendelsohn and Griswold (1967), in recent studies with the *R-S* dimension, also obtained different results for males than females. These studies suggest that the assumption of no sex differences on the *R-S* scale may be rather tenuous, perhaps depending upon the other variables in the investigation.

The finding that males who terminate counseling after three or fewer sessions tend to score lower on the *R-S* scale is subject to at least two interpretations. First, the relationship between adjustment and *R-S* allows for the possibility that many Ss staying in therapy are simply more maladjusted than the early terminators. It was possible to get some information about this by comparing those males in the T group who attended three or less therapy sessions with those attending four or more sessions on those CPI scales which have no more than one item overlapping with the *R-S* scale. The *t* tests (two-tailed) indicate that the two groups do not differ appreciably on the Responsibility, Communality, and Achievement via Independence scales. However, the early terminators show better adjustment on the Flexibility scale ($t = 2.60$, $p < .02$) and on the Psychological-Mindedness scale ($t = 2.04$, $p < .10$). The second possibility is that, since sensitizers tend to approach their problems, they are more likely to pursue their problems in therapy. Future research should investigate the relationship between *R-S* and premature termination while controlling for adjustment. Premature termination should be measured in a way which reliably assesses a client's retreat from psycho-

therapy. Also, if repressors tend to avoid confronting personal problems, future research might profitably explore the relationship between R-S and progress in therapy.

REFERENCES

- BRUNER, J. S., & POSTMAN, L. Emotional selectivity in perception and reaction. *Journal of Personality*, 1947, 16, 69-77.
- BYRNE, D. The repression-sensitization scale: Rationale, reliability, and validity. *Journal of Personality*, 1961, 29, 334-349.
- BYRNE, D. Repression-sensitization as a dimension of personality. In B. A. Maher (Ed.), *Progress in experimental personality research*. New York: Academic Press, 1964.
- BYRNE, D., BARRY, J., & NELSON, D. Relation of the revised repression-sensitization scale to measures of self-description. *Psychological Reports*, 1963, 13, 323-334.
- BYRNE, D., GOLIGHTLY, C., & SHEFFIELD, J. The Repression-Sensitization scale as a measure of adjustment: Relationship with the CPI. *Journal of Consulting Psychology*, 1965, 29, 586-589.
- FEDER, C. Z. Relationship of repression-sensitization to adjustment status, social desirability, and acquiescence response set. *Journal of Consulting Psychology*, 1967, 31, 401-406.
- GOLIN, S., HERRON, E. W., TAKOTA, R., & REINECK, L. Factor analytic study of the manifest anxiety, extraversion, and repression-sensitization scales. *Journal of Consulting Psychology*, 1967, 31, 564-569.
- JOY, V. L. Repression-sensitization and interpersonal behavior. In J. Altrocchi (Chm.), Current research developments with the repression-sensitization personality dimension. Symposium presented at the meeting of the American Psychological Association, Philadelphia, August 1963.
- LEPCOURT, H. M. Repression-sensitization: A measure of the evaluation of emotion expression. *Journal of Consulting Psychology*, 1966, 30, 444-449.
- MENDELSON, G., & GRISWOLD, B. B. Anxiety and repression as predictors of the use of incidental cues in problem solving. *Journal of Personality and Social Psychology*, 1967, 6, 353-359.
- MERBAUM, M., & BADIA, P. Tolerance of repressors and sensitizers to noxious stimulation. *Journal of Abnormal Psychology*, 1967, 72, 349-353.
- TEMPONE, V. J., & LAMB, W. Repression-sensitization and its relation to measures of adjustment and conflict. *Journal of Consulting Psychology*, 1967, 31, 131-136.
- TUCKY, A. W., & GRIGG, A. E. Repression-sensitization as a variable in deviant responding. *Journal of Clinical Psychology*, 1964, 20, 92-93.

(Received March 18, 1968)

SEXUAL ASPECTS OF TAT ADMINISTRATION

JOSEPH MASLING¹ AND SANDRA HARRIS

State University of New York at Buffalo

This study investigated the frequency of administration of four sexual-romantic TAT cards by 20 male and six female graduate student examiners to clinic clients. Male examiners gave more of these cards to female clients than to male clients ($p = .01$) but female examiners did not make such discriminations. The results confirm Schafer's observation that the examiner's voyeuristic needs may be gratified through testing and parallel the findings of Rosenthal that male *Es* behave differently to female *Ss* than to male *Ss*.

While it might be comforting to believe that the psychological examiner (*E*) is a perfectly standardized instrument, never permitting any personal aspect of himself to intrude in the testing situation, evidence collected in recent years shows beyond any doubt that the *E-S* relationship is witness to most of the vicissitudes of any two-person interaction. The personality, needs, and expectations of the psychologist have been shown to influence the administration (Masling, 1959), scoring and interpretation (Haase, 1955; Masling, 1957), and eliciting of responses (Masling, 1965; Turner & Coleman, 1962) of psychological tests.

The importance of the sex of the *E-S* pair has been well described in the psychological experiment (Rosenthal, 1966, 1967), but has not been given adequate attention in the testing situation. Only one aspect of the sexual situation has been investigated—the influence of the sex of the *E* on *Ss*' projective test responses—with mixed results: three studies reported no differences (Alden & Benton, 1951; Garfield, Blek, & Melker, 1952; Holtzman, 1952), one study reported significant differences (Clark, 1952), and one study reported that sexual differences sometimes occurred and sometimes did not (Rabin, Nelson, & Clark, 1954). Perhaps the most penetrating analysis of the sexual aspects of testing has been provided by Schafer (1954) in his discussion of the voyeuristic needs of *E*:

The tester is in the position of a psychological voyeur. He peeps into the interior of many individuals and never once commits himself, as would be

required under normal conditions, to a relationship. . . . Going further, our knowledge of primitive, affect-laden levels of thinking suggests that psychological voyeurism may be unconsciously elaborated as an act of hostile, sexual intrusion. That is to say, even if the tester did not, to begin with, choose testing in part as a more or less sublimated outlet for his infantile voyeuristic inclinations, these inclinations may well seize on the looking-in-secret aspects of testing and thus may invade the professional role [p. 21].

One possible peeping situation arises when *Ss* are confronted with sexual associations suggested by projective test stimuli. Although strangers do not ordinarily discuss sexual matters with each other, particularly with someone of the opposite sex (Benney, Riesman, & Star, 1956), *E* is in a position of forcing an *S* to look at and talk about sexual stimuli.

The selection of a battery of tests is ordinarily determined by the individual needs of the client, the background and experience of the clinician, and the time available for the diagnostic session. Within these broad limits the clinician has a good deal of freedom to choose the testing instruments he will use, and if he uses the TAT he has the option of selecting among 31 cards in the TAT set. The selection of particular TAT cards, therefore, allows *E* to project his own needs into the testing situation. If there is a voyeuristic need in some *Es*, or a need to be sexually forward or provocative, one would expect that there would be a nonrandom use of particular TAT cards, those portraying obvious sexual or romantic themes, as a function of sex and age of *E* and *S*.

The following hypotheses were tested in this study:

¹Requests for reprints should be sent to Joseph M. Masling, Psychological Clinic, State University of New York at Buffalo, Buffalo, New York 14214.

1. When an *E* tests a male *S* and a female *S* he will use more TAT cards with obvious sexual and romantic themes for the cross-sex *S* than the same-sex *S*.

2. The *Ss* in a cross-sex testing situation will be given more TAT cards with obvious sexual and romantic themes than *Ss* in a same-sex testing situation.

METHOD

Examiners

The *Es* were 20 male and six female graduate students in clinical psychology, all at the internship or postinternship level, working for the Psychological Clinic of the State University of New York at Buffalo. The clinic files for the period between February 1958 and November 1963 were inspected and all administrations of the TAT were recorded. Nine of the male *Es* and five of the female *Es* tested both a male and female *S*, thus allowing a separate analysis of the data in which each *E* could serve as his own control.

Subjects

The test *Ss* were clients requesting psychological services from the clinic. They represented the usual range of symptomatology seen at a University clinic, from transitory emotional problems to severe neurotic disturbances. Only the test records of *Ss* who were approximately the same age as the *Es* were used.

TAT Cards

There are only five cards in the TAT set which portray the interaction of a male and female of about the same age, and of these one card, 6 GF, is labeled primarily for use with female *Ss* and was therefore not considered in this study. The cards used were 2, 4, 10, and 13 MF: 2 portrays a young girl holding school books and a young farm worker, naked to the waist; 4 shows a young woman trying to detain a young man from leaving her; 10 presents the faces of a couple, apparently in an embrace; 13 MF shows a woman in bed, nude to the waist, with a young man in the foreground covering his face.

Treatment of the Data

A tally was kept of the number of times cards 2, 4, 10, and 13 MF were administered; the sex of *S* and *E*; and the total number of TAT cards used. Hypothesis 1 was tested using the 14 *Es* who had administered the TAT to both a male and female *S*. Hypothesis 2 was tested using one case from all 26 *Es* in the sample; where *Es* had more than one appropriate case the selection was made on a random basis.

RESULTS

The data, each *E* being used as his own control, are presented in Table 1. In no cir-

TABLE 1

USE OF TAT CARDS 2, 4, 10, AND 13 MF BY
MALE AND FEMALE EXAMINERS WHEN
THEY TEST ONE MALE AND ONE
FEMALE SUBJECT

Examiner	More cards to male <i>Ss</i>	More cards to female <i>Ss</i>	Same num- ber to each
Male	0	5	4
Female	1	1	3

Note.—*N* = 14.

cumstance did a male *E* give more of the sexual-romantic cards to a male *S* than to a female *S*, while four male *Es* gave the same number of cards to the male and female *Ss*. If the tied instances are disregarded, a binomial expansion (Siegel, 1956) for the likelihood that all five nontied events would be found in the male *E*-female *S* cell produced $p = .03$. There was no significant difference in the distribution of the five female *Es* regarding card choice. The data for the male *Es* suggest confirmation of Hypothesis 1 when the tied cases are not considered, but no support was found for the female *Es*.

Hypothesis 2 similarly is supported for the male *Es* but not for the female *Es*. Table 2 shows the distribution of the four sexual-romantic cards to male and female *Ss* by male and female *Es*. Male *Es* clearly gave more of these cards to female *Ss* than to male *Ss* ($p = .01$, Fisher exact test, Siegel, 1956), whereas the female *Es* treated male and female *Ss* alike.

To test for the possibility that this distribution of TAT cards was primarily a function of the larger number of cards used by the male *Es*, correlations were computed between

TABLE 2

USE OF TAT CARDS 2, 4, 10, AND 13 MF

Cards	By male examiners*		By female examiners	
	Given to male <i>Ss</i>	Given to female <i>Ss</i>	Given to male <i>Ss</i>	Given to female <i>Ss</i>
3 or 4	1	9	1	1
1 or 2	7	3	2	2

Note.—For male examiners, *N* = 20; for female examiners, *N* = 6.

* $p = .01$ (Fisher exact test).

the total number of cards administered and number of sexual-romantic cards used. For the male *Es*, the correlations were $-.20$ for the male *Ss* and $.16$ for the female *Ss*; for the female *Es*, the correlations were $.00$ for the male *Ss* and $.11$ for the female *Ss*. None of these correlations was significantly greater than zero. As a further test, Hypothesis 2 was evaluated by converting the number of sexual-romantic cards each *E* administered into a proportion of all TAT cards he used; this measure produced results identical to that reported in Table 1: Fisher exact test = $.01$ for the male *Es* and no significant differences for the female *Es*.

Inspection of the use of the four sexual-romantic cards showed that the significant difference in Table 2 was caused by male *Es* giving Cards 4 and 10 more frequently to the female *Ss* than male *Ss*. The Fisher exact test computed on these data showed $p = .06$ for Card 4 and $p = .05$ for Card 10. While the male *Es* varied in the use of these cards as a function of the sex of *S*, the female *Es* showed a different pattern: with one exception, they either gave the card to all *Ss* or they gave them to none. Evidently, the female *Ss* were concerned almost exclusively with the stimulus quality of the TAT cards, while the male *Es* seemed to have one eye on the card and the other on the sex of the *S*. Card 13 MF, the most obviously sexual of all the TAT set, did not discriminate because it was so widely used by all *Es*.

If the male *E*-female *S* relationship contains some element of sexual play, there would be a temptation for the male *E* to prolong the session by administering more TAT cards to female *Ss* than to male *Ss*. Table 3 shows a

distribution of total number of TAT cards used by male and female *Es*. As was found in the previous analyses, the female *Es* treated male *Ss* and female *Ss* equally, but there was a nonsignificant trend for the male *Es* to give more cards to female *Ss* than to males ($p = .16$, Fisher exact test, Siegel, 1956).

DISCUSSION

In general, this study showed that female clinicians were extremely task-oriented, while the males were distracted somewhat by more personal considerations. The observation of Schafer (1954) that *Es* may use the testing situation to gratify voyeuristic needs was confirmed for males. The report of Kinsey, Pomeroy, Martin, and Gebhard (1953) that voyeurism is almost exclusively a male activity was also supported here. The results also parallel the conclusion reached by Rosenthal (1967) based on his investigation of the psychological experiment that male *Es* pay more attention to the sex of *S* than do female *Es*.

Unfortunately, the use of internship level graduate student *Es* limits the confidence with which these findings can be generalized to older clinicians, although the sample size is much larger than those commonly employed in experiments in this area. Mature, experienced clinicians may never behave like this, or scarcely ever. But apart from the use of graduate students, this study sampled a genuine clinical situation—real clients were seen in a real clinic for important problems. The results suggest that the observation of some clients (and others) that psychologists (male) think about sex a great deal may not be totally in error.

TABLE 3

TOTAL NUMBER OF TAT CARDS ADMINISTERED

Cards	By male examiners*		By female examiners	
	Given to male <i>Ss</i>	Given to female <i>Ss</i>	Given to male <i>Ss</i>	Given to female <i>Ss</i>
11 or more	2	7	0	0
0 to 10	6	5	3	3

Note.—For male examiners, $N = 20$; for female examiners, $N = 6$.

* $p = .16$ (Fisher exact test).

REFERENCES

- ALDEN, P., & BENTON, A. L. Relationship of sex of examiner to incidence of Rorschach responses with sexual content. *Journal of Projective Techniques*, 1951, 15, 230-234.
- BENNEY, M., RIESMAN, D., & STAR, S. A. Age and sex in the interview. *American Journal of Sociology*, 1956, 62, 143-152.
- CLARK, R. A. The projective measurement of experimentally induced levels of sexual motivation. *Journal of Experimental Psychology*, 1952, 44, 391-399.
- GARFIELD, S., BLEK, L., & MELKER, F. The influence of method of administration and sex differences on

- selected aspects of TAT stories. *Journal of Consulting Psychology*, 1952, 16, 140-145.
- HAASE, W. Rorschach diagnosis, socio-economic class, and examiner bias. Unpublished doctoral dissertation, New York University, 1955.
- HOLTZMAN, W. H. The examiner as a variable in the Draw-A-Person Test. *Journal of Consulting Psychology*, 1952, 16, 145-148.
- KINSEY, A. C., POMEROY, W. B., MARTIN, C. E., & GEBHARD, P. H. *Sexual behavior in the human female*. Philadelphia: Saunders, 1953.
- MASLING, J. The effects of warm and cold interaction on the interpretation of a projective protocol. *Journal of Projective Techniques*, 1957, 21, 377-383.
- MASLING, J. The effects of warm and cold interaction on the administration and scoring of an intelligence test. *Journal of Consulting Psychology*, 1959, 23, 336-341.
- MASLING, J. Differential indoctrination of examiners and Rorschach responses. *Journal of Consulting Psychology*, 1965, 29, 198-201.
- RABIN, A., NELSON, W., & CLARK, M. Rorschach content as a function of perceptual experience and sex of the examiner. *Journal of Clinical Psychology*, 1954, 10, 188-190.
- ROSENTHAL, R. *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts, 1966.
- ROSENTHAL, R. Covert communication in the psychological experiment. *Psychological Bulletin*, 1967, 67, 356-367.
- SCHAFER, R. *Psychoanalytic interpretation in Rorschach testing*. New York: Grune & Stratton, 1954.
- SIEGEL, S. *Nonparametric statistics*. New York: McGraw-Hill, 1956.
- TURNER, G. C., & COLEMAN, J. Examiner influence on Thematic Apperception responses. *Journal of Projective Techniques*, 1962, 26, 478-486.

(Received March 26, 1968)

ANALITY AND TIME

TUPPER F. PETTIT¹

New York University

A Time Scale (TS) of 40 items was constructed to ascertain the ways in which time may have special meaning and/or may serve to organize an individual's experience. The TS and other scales were administered to 91 Ss. The TS had significant correlations of .643 with the Composite Analilty Scale, of .507 with the Grygier Analilty Scale, and of $-.316$ with the Grygier Spontaneity Scale. All correlations were in the predicted direction. The results were interpreted as in accord with the psychoanalytic formulation of the anal character, which makes it possible to relate the intensity of certain character traits to an individual's estimation of the importance of time.

Most investigators (e.g., Bell, 1966; du Preez, 1967) have concentrated on time as a physical dimension, studying such variables as estimation of time intervals, speed of tapping, and the like. In these experiments it seems to be generally recognized that there are individual differences which give rise to differences in time estimation, but these dimensions have yet to be specified in any systematic way.

A fruitful approach to understanding the importance of time as a psychological dimension may be to relate an individual's sense of time to his relatively enduring character traits. In a recent book, Cohen (1967) mentioned briefly the value which can accrue to time, some people being parsimonious with it and others being profligate with it. In these or other circumstances time may have a symbolic meaning. Such an explanation has been advanced in psychoanalysis with reference to the anal character, although Cohen dismisses this as "interesting speculation" which is difficult to verify.

The Anal Character

In "Character and Anal Erotism" Freud (1962) first revealed his observation that a certain cluster of character traits are frequently found in one individual. These traits are orderliness, parsimony, and obstinacy. He observed that the appearance of this triad

of character traits in an individual coincided with the disappearance of his anal erotism.

Freud conceived of the dynamics that give rise to such a triad in the following way. If a conflict occurs between the mother and child over toilet training, competition arises between the two. The only weapon the child has at his disposal against the all-powerful mother is his feces which both he and his mother regard as having value. Given these circumstances, the accumulation of feces provides the infant with the dual satisfaction of a pleasurable tension and defiance of the mother. The control of his sphincter, both holding on and letting go, is the simultaneous control over his pleasurable tension states and an exertion of autonomy. The situation is exacerbated as the mother comes to place more demands on her child for both production and control.

In this case, an inordinate amount of libidinal energy becomes fixated at the anal stage of development. But since there is little satisfaction to be gained for anal erotism in our culture, the child develops a reaction formation against his anal erotic desires and the anal products during the latency period. The reaction formation affects both feelings about anal activities and character traits, and the two are inextricably linked. The original pleasure and satisfaction associated with anal activities are replaced by feelings of shame and disgust. Consequently, the triad of character traits—orderliness, parsimony, and obstinacy—become prominent in the personality.

Since the anal character was first postu-

¹ The author wishes to thank Irving Sarnoff and M. E. Tresselt for their assistance in the preparation of this paper. Requests for reprints should be sent to Tupper F. Pettit, 100 Bleeker Street, No. 14C, New York, New York 10012.

ated, many psychoanalytic writers have substantiated Freud's notions with idiographic data. An aspect which Freud originally commented on, and which has received a good deal of attention, is the anal character's concept of money. To him, money is a symbolic representation of the valued feces and, as such, is dealt with in the same irrational way: to be accumulated and hoarded and occasionally be spent profligately.

Another aspect which has received somewhat less attention is the meaning that time has for the anal character. In many ways time takes on the same irrational, symbolic meaning that money does ("Time is money!"). Fenichel (1946) wrote that in compulsion neuroses the patient's relation to time, like his relation to money, is "an arena wherein his instinctual conflicts are fought out [p. 282]." Abraham (1953) noted that "many neurotics are continually worrying over waste of time. . . . The patients often save time on a small scale and waste it on a great one [p. 384]."

Fraisse (1963) made a statement clearly reminiscent of Freud's characterization of the dual pleasure a child attains from the increased bowel tension and control over reduction of that tension: a person who lacks punctuality may "postpone an activity in order to create a tension which will be a source of satisfaction [p. 290]."

In what way would the traits which are under consideration make themselves manifest in behavior, particularly in regard to time? Using Freud's triad, it is expected that a person with anal character traits would try to *order* his experiences so that they are organized, well delineated, and sharply focused, as opposed to being diffuse and undifferentiated. Time can become very important in ordering experiences by providing a beginning and an end, and in this way becomes more figure than ground. Since the sharpness of and control over his experiences are important, the anal character should dislike spontaneous experiences and surprises, which are, after all, a matter of timing.

In addition, the anal character should be *parsimonious* in the way he uses time and *obstinate* about giving any of it up, that is, wasting it. Jones (1961) observed that

the concept of time is, because of the sense of value attaching to it, an unconscious equivalent of excretory product. . . . People of this type are particularly sensitive about their time being taken up against their will, and they insist in every way on being master of their own time [p. 420].

Schlesinger (1963) did a factor analysis on 154 items which had been used in various anality questionnaires. She extracted 12 factors. These factors and the resulting intercorrelations of factor loadings are in agreement with and tend to confirm theoretical expectations and clinical experience. The 12 factors are as follows:

- I: Responsibility in dealing with others
- II: Regularity and meticulousness
- III: Retentiveness as a style of life
- IV: Obstinacy
- V: Rigidity or intellectual curiosity
- VI: Frugality
- VII: Concern about dirt and contamination
- VIII: Orderliness
- IX: Self-righteous hostility and competitiveness
- X: Anxiety over possible loss of control
- XI: Sensitivity to smells
- XII: Retentiveness in relation to possessions

Rosenwald, Mendelsohn, Fontana, and Portz (1966) took poor performance on a hand-eye coordination task in an odorous, dirty medium reminiscent of fecal matter as an indication of ineffective defenses against anal impulses. They also found significant correlations between ineffective defenses against anal stimuli and other traits thought characteristic of anal characters (namely, indecisiveness and selective sensitivity for anal concepts).

The purpose of this investigation was twofold: to develop a questionnaire which gets at the meaning that time has for individuals and the importance of time as a factor in controlling experiences; and to test the psychoanalytically derived hypothesis that the intensity of an individual's anal character traits is positively related to his estimation of the importance of time.

METHOD

Time Scale (TS)

A time questionnaire was constructed which consisted initially of 76 items. Items were chosen which seemed in keeping with the theoretical formulation previously developed. Several questions were taken directly or in modified form from an 85-item questionnaire developed by Loehlin (1959). The final TS consisted of such items as: I feel that I am more conscientious about being on time than most people I know; I often don't know what the date is; It doesn't bother me if the teacher is late to class; and Time is money.

Composite Anality Scale (CAS)

Two anality questionnaires were used, neither of which contained items having explicit reference to time. The first was developed for this investigation from the study by Schlesinger. The items were selected simply by picking the items which loaded each of her 12 factors the most heavily. This resulted in the CAS, which consisted of 66 items, such as: You accumulate a great many things because you rarely throw anything away; When I see a messy room I have a strong impulse to straighten it out; I use more toilet paper than is necessary; and I feel one of the most embarrassing things people can do is break wind.

Grygier Anality Scale (GAS)

The second anality questionnaire consisted of four subscales (Ad, Ah, Ac, and Aa) containing 41 items taken from the Dynamic Personality Inventory developed by Grygier (1956). The entire inventory consists of 325 items and was constructed within the theoretical framework of psychoanalysis. There are many scales for measuring patterns of psychosexual development, masculine and feminine identification, etc. The four subscales used to form the GAS were Ad, which, according to Grygier, measures "attention to details; orderliness, conscientiousness and perfectionism"; Ah, which measures "hoarding behavior, anxious possessiveness, and stubborn, clinging persistence"; Ac, which measures "conservatism, rigidity, tendency to stick to routine"; and Aa, which measures "submissiveness to authority and order."

TABLE 1

INTERCORRELATIONS FOR ALL SUBJECTS

Scale	Composite Anality Scale	Grygier Anality Scale	Grygier Spontaneity Scale	Sex
Time	.643**	.507**	-.316*	.052
Composite Anality		.570**	-.201	.113
Grygier Anality			-.405**	-.052
Grygier Spontaneity				-.186

Note.—*N* = 91.* *p* < .01.** *p* < .001.

The Grygier scales are administered by asking *S* to indicate whether he "likes" or "dislikes" a certain thing. Some of the questions included in this study were whether *S* likes or dislikes: Perfect order and symmetry (Ad); Hanging onto old clothes you never wear (Ah); Keeping out of trouble as much as you can (Ac); and Patriotism as the highest ideal (Aa). Several recent studies (Bishop, 1967; Gordon, 1966, 1967; Rapaport, 1963) have yielded evidence in support of the validity of Grygier's anality scales.

Grygier Spontaneity Scale (GSS)

The final scale used in this study was the Grygier Of Scale, which measures "impulsiveness, changeability, spontaneity, speed of reaction, emotional expressiveness." It is this type of experience and emotional expression that the anal character guards against by relying heavily on the time dimension to structure his life, since it is often a loss of the sense of time that allows one to be surprised or to react spontaneously. Thus, high scores on the time and anality scales should go hand in hand with low scores on the GSS.

The GSS consisted of 12 items, three of them appearing simultaneously in the GAS. The scale included questions regarding whether *S* liked or disliked: Smoking; Acting on the spur of the moment; and Spending money lavishly.

These four questionnaires were administered to 91 voluntary, undergraduate psychology students in the following order: TS, GAS and GSS combined, and CAS. Thirty-seven *Ss* were males, and 54 were females.

RESULTS

The TS was refined by dropping those items with the lowest item-test correlation. This was done in four steps, with the item-test correlations for the remaining items being recomputed between each step. All statistics reported in this paper are for the final, 40-item version of the TS. The reliability for it, as calculated by Kuder-Richardson Formula 20, was .833.

The reliabilities of the anality scales were calculated using the split-half method. The reliability of the CAS was .811 and of the GAS .682.

Table 1 shows the intercorrelations of the scales and sex for all *Ss*. The TS correlated .643 and .507, respectively, with the CAS² and GAS; both correlations are highly significant. The TS had a significant correlation of $-.316$

² Subsequently, slightly revised 40-item versions of the TS and CAS have been administered to 96 undergraduate psychology students as part of another study. The correlation obtained between the two measures was .611.

TABLE 2
INTERCORRELATIONS FOR MALE SUBJECTS AND FEMALE SUBJECTS

Scale	Composite Anality Scale		Grygier Anality Scale		Grygier Spontaneity Scale	
	Male	Female	Male	Female	Male	Female
Time	.667**	.621**	.506*	.520**	-.270	-.344*
Composite Anality			.650**	.542**	-.164	-.202
Grygier Anality					-.169	-.577**

Note.—For males, $N = 37$; for females, $N = 54$.

* $p < .01$.

** $p < .001$.

with the GSS. Sex was not a significant factor on any of the scales for the pooled Ss.

To determine if males and females scored differently on the scales, correlations were calculated for the males as a group and for the females as a group on the different scales. The data are presented in Table 2. The differences between corresponding correlations are in each case nonsignificant except for the correlations between the GAS and GSS. This correlation is $-.169$ for the males and $-.570$ for the females, and the difference between them is significant at the .05 level.

DISCUSSION

This study gives considerable support to the theoretical relation between time and anality. The correlation between the TS and the CAS is unexpectedly high. Based on this sample of 91 Ss, the .99 confidence limits for the population r are .35 and .72. Thus it seems safe to conclude that there is a sizable relationship between these two dimensions and that time does have special meaning for people with anal character traits. The next step is to establish more precisely, through experimental means, the nature of that relationship and the centrality of its importance to such individuals.

An objection which may be raised is that social desirability rather than anal character traits is the mediating variable between the time and anality scales. While the issue cannot be dealt with empirically here, several factors would seem to militate against such an explanation. One is that many items refer to acts of behavior in which the socially desirable direction of endorsement would be difficult to

determine; therefore, what an individual really does in a given situation is liable to be the overriding influence in the response. The second factor is that several items which were clearly socially undesirable correlated very highly, while other items which could be easily interpreted in a socially desirable way had very low item-test correlations.

Of the two anality scales, the CAS seems to be the superior one. Its reliability is much higher and its content seems more directly related to the dimensions of anality. The reliability of the GAS is so low as to make fine discriminations between Ss questionable. The things which seem to be working against it are its comparatively short length, the unfamiliarity of the "Like-Dislike" format, and the outdated quality of many items—at least for many contemporary college students. Examples of such questions are whether *S* likes: Absolute respect for one's parents, without reservation; Conscientious objectors; Wearing a uniform; and Keeping a large stock of canned goods in your cupboard.

The GSS yielded correlations with the other scales which seem equivocal. The significant correlation between the GSS and TS was as expected, but the correlation of $-.201$ between the GSS and CAS is too low to enable firm conclusions to be drawn. The relatively high correlation of $-.405$ between the GSS and GAS is undoubtedly spuriously high because of three items which appear in both scales, and this constitutes 25% of the GSS. The results only suggest that the theoretical relationship between time, anality, and spontaneity is valid.

The significant difference found between the

correlations of the GAS and the GSS for males and females is likely an artifact of the Grygier scales, since it is difficult to explain theoretically, and it is not confirmed in the correlations between the CAS and GSS.

It was the psychoanalytic formulations of the variables under consideration which led to this study, and in this way the heuristic value of psychoanalytic theory is demonstrated. The two specific and important formulations which derive from the theory are (a) the way in which enduring character traits may be the result of the handling of early sexual impulses, and (b) the way in which time may take on symbolic meaning.

An alternate explanation which could be advanced to explain the relationships found is a cultural one. That is, traits which are considered "anal" may not be the result of a defense against anal impulses, but rather internalized cultural values. In fact both factors are probably present and it would take a combination of longitudinal and experimental studies to determine their complex relation. A serious complication in such studies is that many of the relevant cultural values are passed on to the child at the time of toilet training. However, the one distinguishing feature between the two explanations—one that could be investigated experimentally—is that in the former time would take on *symbolic* value, but it would not in the latter.

The present study would seem to show the value of relating the dimension of time to enduring personality traits, whether or not one wishes to accept the specific theoretical formulations which are central to this paper.

REFERENCES

- ABRAHAM, K. Contributions to the theory of the anal character (1921). In, *Selected papers of Karl Abraham, M.D.* New York: Basic Books, 1953.

- BELL, C. R. Personality and judgments of temporal intervals. *British Journal of Psychology*, 1966, 57, 155-159.
- BISHOP, F. V. The anal character: A rebel in the dissonance family. *Journal of Personality and Social Psychology*, 1967, 6, 23-36.
- COHEN, J. *Psychological time in health and disease*. Springfield, Ill.: Charles C Thomas, 1967.
- DU PREEZE, P. Production of time intervals after short periods of delay. *Journal of General Psychology*, 1967, 67, 59-71.
- FENICHEL, O. *The psychoanalytic theory of neurosis*. London: Routledge & Kegan Paul, 1946.
- FRAISSE, P. *The psychology of time*. New York: Harper & Row, 1963.
- FREUD, S. Character and anal erotism. In, *Collected papers*. Vol. 2. London: Hogarth Press, 1962.
- GORDON, C. M. Some effects of information, situation, and personality on decision-making in a clinical setting. *Journal of Consulting Psychology*, 1966, 30, 219-224.
- GORDON, C. M. Some effects of clinician and patient personality on decision-making in a clinical setting. *Journal of Consulting Psychology*, 1967, 31, 477-480.
- GRYGIER, T. C. *Dynamic Personality Inventory*. Sulken, England: Banstead Hospital Management Committee, 1956.
- JONES, E. Anal erotic character traits (1918). In, *Papers on psychoanalysis*. Boston: Beacon Press, 1961.
- LOEBLIN, J. C. The influence of different activities on the apparent length of time. *Psychological Monographs*, 1959, 73(4, Whole No. 474).
- RAPAPORT, C. Character, anxiety and social affiliation. Unpublished doctoral dissertation, New York University, 1963.
- ROSENWALD, C., MENDELSON, G. A., FONTANA A., & PORTZ, A. T. An action test of hypotheses concerning the anal personality. *Journal of Abnormal Personality*, 1966, 71, 304-309.
- SCHLESINGER, V. J. Anal personality traits and occupational choice: A study of accountants, chemical engineers, and educational psychologists. Unpublished doctoral dissertation, University of Michigan, 1963.

(Received March 26, 1968)

VALIDITY OF INFORMANTS' RATINGS OF THE BEHAVIOR AND SYMPTOMS OF PSYCHIATRIC PATIENTS

NORRIS D. VESTRE AND ROBERT ZIMMERMANN¹

University of Minnesota

Psychiatric patients were rated with respect to symptoms and behavior by their nearest relative and also by a nurse after admission to a psychiatric hospital. Data on 138 consecutive admissions were analyzed by computing product-moment correlations between 12 family rating scale clusters and four ward behavior factors and by canonical correlations between the two sets of data. The canonical analysis indicated there were two underlying factors accounting for the intercorrelations among the two sets of ratings. The results of both analyses were seen as providing support for the validity of relatives' ratings of psychiatric symptoms and behavior.

One of the critical problems in mental health research is the need for reliable and valid procedures of assessing psychiatric condition and change in psychiatric condition. This is particularly true in the case of patients living in the community where opportunities for observation by the psychiatric staff are severely limited. Adequate procedures for evaluating the symptoms and behavior of such patients are needed.

Recent changes in the philosophies of treatment as well as in the methods of treatment have contributed to shorter hospital stays. These shorter stays have often been accompanied by an increase in the readmission rate. Moreover, there has been an increased tendency for hospitals to use disposition procedures such as provisional discharge or trial visit as opposed to a regular or direct discharge. A practical and efficient procedure of following these released patients on a long-term basis could prove useful. Any research concerned with the long-term effectiveness of treatment has been forced to either obtain generous financial support or rely on such practical criteria such as readmission rate or amount of time the patient remains out of the hospital. Although these criteria are easily obtained, they are also poorly understood because other factors, in addition to the nature and severity of the patients' symptoms, would seem to influence them. Consequently,

it would be highly desirable to have an easily obtained, yet meaningful, procedure for assessing the behavior and symptoms of the patient in the community.

One obvious procedure of assessing the psychiatric status of patients in the community would be to have a relative of the patient provide the data. Katz and Lysterly (1963) developed a standard set of items covering various symptom and behavior areas. The fact that relatives' ratings of symptoms and behaviors formed identifiable clusters of psychopathology, when cluster-analyzed, probably gives some evidence of validity of their ratings. Katz and Lysterly (1963), however, have provided some additional data on the discriminative validity of their scale. Also, Ellsworth, Arthur, Kroeker, and Childers (1967) offered some data to indicate that the ratings of relatives were reliable and valid. Small, Small, and Gonzalez (1965) concluded that relatives, as informants on the prehospitalization symptoms of patients, provided a more complete picture than did the patients themselves.

The purpose of the present study was to explore further the validity of relatives' ratings of patient behavior by examining the correspondence between the ratings made by the relatives and those made by the staff. In addition, this study was intended to provide data bearing specifically on the validity of the Katz Adjustment Scales (KAS; Katz & Lysterly, 1963) as a device for assessing the prehospitalization symptoms and behavior of psychiatric patients.

¹The authors acknowledge the assistance of Dean J. Clyde in the computation of the canonical correlations. Requests for reprints should be sent to Norris D. Vestre, Department of Psychology, Indiana State University, Terre Haute, Indiana 47809.

METHOD

Sample and Procedure

Consecutively admitted patients with functional psychiatric disorders were selected for study. Within 3 days of admission to the psychiatric ward of a university hospital, each patient's nearest relative was mailed a KAS. Of all the KAS mailed, 83% were returned and usable. The Psychotic Reaction Profile (PRP; Lorr, 1961) was completed by the nurse for each patient within one week of admission to the hospital.

Complete data, that is, a KAS plus a PRP, were obtained on 159 patients. For 138 of these, 87%, the rater was either the spouse or parent(s). KAS data for the remaining 21 cases were obtained from such a wide variety of raters, including friends, siblings, children, and more distant relatives that it was decided to exclude these from the major analyses of the data.

The 138 patients included in the analyses ranged in age from 16 to 65 with a median age of 26 years. Females made up 69.1% of the sample. The diagnostic breakdown of the sample was as follows: schizophrenics, 49%; other psychotics, 16%; neurotics, 16%; and others (including the personality pattern and trait disturbances, the sociopathic personality disturbances and the transient situational personality disorders), 19%.

Katz Adjustment Scale

The KAS (Form R1) is designed for use by a relative of a patient in rating the patient's symptoms

and social behavior. The 127 items, written to cover typical aspects of psychopathology, are to be judged on a 4-point scale of frequency of occurrence. The KAS can be scored according to a set of 12 item-clusters obtained on the basis of a cluster analysis of the 127-item set (Katz & Lyerly, 1963).

Psychotic Reaction Profile

The PRP (Lorr, 1961) consists of 85 statements describing behavior commonly observed in hospitalized psychotics and is to be completed by a nurse or aid on the basis of observations of the patient. The PRP yields scores in four relatively independent areas of psychopathology, namely, Withdrawal, Thinking Disorganization, Paranoid Belligerence, and Agitated Depression. The validity of the PRP as a measure of severity of psychopathology has been demonstrated by specific validation studies (Lorr, 1961; Vestre, 1966) and by its sensitivity to change in psychiatric condition in response to therapy (Casey, Hollister, Klett, Lasky, & Caffey, 1961).

Since the data were collected over a period of 18 months on three different wards, it was necessary to involve 30 different nurses to obtain ratings. Over two-thirds of the ratings, however, were completed by nine of these nurses. No assessment of interrater reliability was obtained within the present study, but previous data by Lorr (1961) and by Vestre (1966) indicated reliabilities ranging from .65 to .90 and .58 to .92, respectively, for the separate scales. In both studies the Agitated Depression scale yielded the lowest coefficient. Also, the fact that this group of nurses had considerable experience in rating patients for research purposes, using the PRP and similar scales, provides some confidence in the reliability of their ratings.

TABLE 1
KAS-PRP INTERCORRELATIONS

KAS clusters	PRP scales			
	W	TD	PB	AD
Belligerence	-.07	-.07	.05	-.02
Verbal Expansiveness	-.07	.09	.09	-.01
Negativism	-.02	-.05	.03	-.11
Helplessness	.14	.01	-.01	.28**
Suspiciousness	-.04	.06	.00	.12
Anxiety	.05	.07	.01	.27**
Withdrawal-Retardation	.32**	.07	-.10	.15
General Psychopathology	.08	.00	.08	.07
Nervousness	.04	-.07	.08	.19*
Confusion	.10	.09	.11	.13
Bizarreness	.13	.28**	.02	.23**
Hyperactivity	.01	-.02	.05	.14

Note.—For the PRP, W = Withdrawal, TD = Thinking Disorganization, PB = Paranoid Belligerence, AD = Agitated Depression. Decimals omitted. $df = 136$.

* $p < .05$.

** $p < .01$.

RESULTS

Product-moment correlations between the PRP and the KAS are presented in Table 1. Five of the 12 KAS variables correlate significantly with one or more of the PRP scales. It can also be seen from Table 1 that three of the four PRP variables correlate significantly with at least one of the KAS clusters.

A qualitative matching of the factors comprising the two sets, on the basis of descriptions furnished by Lorr (1961) and by Katz and Lyerly (1963), suggested that certain factors should be somewhat parallel. In Table 2 scales with similar item content are paired. For three of four PRP scales, at least one of these pairings is statistically significant at the .01 level and in the expected direction. For the remaining scale, Paranoid Belligerence, none of the pairings approaches statistical significance. With the exception of the pairings with Paranoid Belligerence, the only

TABLE 2

CORRELATIONS BETWEEN PAIRS OF PRP AND KAS SCALES MATCHED FOR SIMILAR ITEM CONTENT, ACCORDING TO TYPE OF FAMILY RATER

PRP scales	KAS clusters	Total group	Rated by spouse	Rated by parent
Withdrawal	Withdrawal-Retardation	32**	36**	29*
Thinking Disorganization	Confusion	09	17	01
	Bizarreness	28**	42**	12
Paranoid Belligerence	Belligerence	05	00	09
	Verbal Expansiveness	09	03	18
	Negativism	03	07	02
	Suspiciousness	00	00	02
	General Psychopathology	08	11	07
Agitated Depression	Helplessness	28**	23	24*
	Anxiety	27**	16	32**
	Nervousness	19*	14	19

Note.—Decimals omitted. For total group, $df = 136$; for remaining two groups, $df = 67$.

* $p < .05$.
 ** $p < .01$.

matched pair of scales not significantly correlated is that of Thinking Disorganization and Confusion. The KAS cluster, Hyperactivity, could not be reasonably matched with any of the four PRP scales.

Of the 48 intercorrelations in Table 1, five reach significance at the .01 level, with one additional correlation significant at the .05 level. Five of these significant correlations occur between variables matched for similar content.

Even though the qualitative comparison of factors between the two rating devices confirmed most of the expected relationships, multivariate procedures were used to determine the degree of confidence which could be placed in the overall hypothesis of an underlying linear relationship between the two sets of variables. Canonical correlations (Cooley & Lohnes, 1962) were computed between the PRP scales and the KAS clusters for the purpose of determining the number of in-

dependent factors influencing both sets of variables.

The canonical correlations and the results of significance tests are shown in Table 3. As Cooley and Lohnes (1962, p. 37) indicate, the chi-square approximation for the distribution of lambda can be used to test the hypothesis that the two sets of variables are unrelated. The two significant canonical correlations obtained indicate that there are two factors which underlie the intercorrelations among the two sets of ratings. The characteristics of these two factors can be inferred from an inspection of the weights (contributions) of the individual variables. Table 4 shows the weights associated with the individual variables for each of the two significant canonical factors.

Latent Root 1 is defined primarily by high positive loadings on KAS Bizarreness and PRP Thinking Disorganization and by high negative loadings on KAS General Psycho-

TABLE 3

CANONICAL CORRELATIONS OF PRP AND KAS VARIABLES

Latent root	Canonical correlations	Wilks lambda	Chi-square	df	$p <$
1	.48	.515	83.4	48	.002
2	.41	.667	50.8	33	.028
3	.39	.804	27.4	20	.129
4	.23	.945	7.1	9	.626

Note.— df for chi-square for testing Latent Root 1 is p PRP scales $\times q$ KAS clusters, where p is 4 and q is 12. Also, with r roots removed, chi-square is distributed with $df = (p-r)(q-r)$.

TABLE 4

WEIGHTS ASSOCIATED WITH LATENT ROOTS 1 AND 2

Variables	Root 1 weights	Root 2 weights
KAS		
Belligerence	.04	-.19
Verbal Expansiveness	.87	.92
Negativism	-.48	-.96
Helplessness	.48	-.54
Suspiciousness	.41	1.00
Anxiety	-.25	.38
Withdrawal-Retardation	.24	-1.74
General Psychopathology	-1.40	.61
Nervousness	-.14	-.16
Confusion	-.29	.26
Bizarreness	1.54	-.75
Hyperactivity	-.14	.08
PRP		
Withdrawal	-.36	-.91
Thinking Disorganization	.75	.17
Paranoid Belligerence	-.38	.36
Agitated Depression	.40	.14

pathology. Thus, the loadings associated with Latent Root 1 seem to define a dimension reflecting behavior and symptoms of schizophrenic disorganization.

Latent Root 2 is characterized by high negative loadings on PRP Withdrawal and KAS Withdrawal-Retardation. These weights would suggest that this dimension clearly reflects degree of interest in, or responsiveness to, the environment and activity level.

DISCUSSION

In general, the results suggest that there are two common factors underlying the nurses' and the relatives' ratings of symptoms. The first factor clearly reflects schizophrenic disorganization or bizarreness; the other, a dimension of withdrawal. It might be noted that there was some suggestion of a third factor emerging in the canonical analysis (see Table 3). Examination of the associated weights suggested a dimension reflecting depression and agitation in that PRP Agitated Depression and KAS Anxiety and Helplessness obtained the highest loadings. Although the results were somewhat inconsistent, and, as can be seen in Table 3, did not reach significance, the zero-order correlations of the PRP Agitated Depression scale (Table 2) are

consistent with such a factor. Thus, the results would seem to indicate that there is, indeed, overlap between symptoms and behavior as viewed by relatives of patients and as viewed by the professional staff.

There are certain facts which should be borne in mind in interpreting the magnitude of the KAS-PRP intercorrelations in Table 1. Most obvious, of course, are the differences in the nature of the raters (nurses vs. relatives of patients) which presumably could result in differences in objectivity or bias in observing and reporting behavior and symptoms and also differences in knowledge of psychiatric symptoms. The fact that the KAS items are written in nontechnical language and refer to rather specific aspects of observable behavior should reduce both of these potential problems.

A more important fact to be borne in mind in interpreting the magnitude of the Table 1 intercorrelations, however, relates to differences in time at which and setting in which the two sets of ratings were obtained. Relatives based their ratings on the behavior of the patient during the days just prior to admission; nurses based their ratings on the behavior of the patient during the first week of hospitalization. Perhaps even more important than the 1- to 2-week time differential, would be the differences in the setting or situation. Patients show wide differences in how they respond to being hospitalized. Being admitted to a psychiatric ward tends to give some patients a feeling of comfort and security, for example, a patient who is disturbed and frightened by suicidal impulses. For other patients, admission to the hospital is a disturbing or upsetting experience which leads to further disorganization of behavior. Consequently, the time differential between the two sets of ratings and the different settings in which the observations were made can be seen as factors which would reduce the relationships between the two sets of observers. In fact, one explanation for the absence of relationships involving PRP Paranoid Belligerence might be along these lines. Of the four dimensions (scales) of psychopathology measured by the PRP, it would seem that Paranoid Belligerence, that is, hostility, irritability, resistiveness, and suspiciousness, would

be the most dependent upon the situation or environment.

Table 2 presents the correlations by type of family rater for those scales matched for similar item content. The two sets are reasonably similar and the differences could easily result from chance variation. The largest difference is in the low correlations between the two measures for Thinking Disorganization in the group rated by parents, as opposed to the group rated by the spouse. The KAS-PRP correlations for the 21 patients rated by persons other than parents or spouse were markedly different, but the small size of the group makes it difficult to be certain that this is a real difference. This difference in the correlation matrix and the disparate nature of the raters in this small group are the reasons for excluding these cases from the canonical correlations.

The finding that there is a relationship between the ratings of symptoms and behavior made by relatives and those by the psychiatric staff suggests that relatives might be a useful source of such information for research purposes. While canonical correlations tend to decrease on replication, due to the fact that they capitalize on chance combinations, the general coherence of the variables, along with the significant and meaningful zero order correlations, suggest that a valid relationship exists between these two sets of variables. Despite the possibility that the relatives may lack objectivity in observing and reporting behavior and symptoms, the results of this study suggest their ratings are valid descriptions of the nature and severity of such behavior and symptoms.

While the magnitude of the correlations are not large enough to allow clinical interpretation or individual prediction, they are large enough to indicate the KAS is a useful research tool. Moreover, the present correlations represent the minimal validity of a relatively new and undeveloped instrument. Efforts could be directed toward improving the validity of the instrument. First, the clusters themselves may need refinement and a factor

analysis of the items using a larger sample and more objective methods of rotation may yield more homogeneous factors. In addition, validity could be increased without increasing the length of the total instrument by dropping items not scored on any cluster and adding items to strengthen those clusters containing only four or five items.

An entirely different approach to improving validity might involve a study of variables related to valid ratings, that is, a study of moderator variables. The most obvious and accessible of these are associated with the relationship of the rater to the patient, not only its objective aspects, but its subjective components as well. Such a line of research should yield refinements which will make the KAS an even more useful instrument in an area where too little is available at the present time.

REFERENCES

- CASEY, J. F., HOLLISTER, L. E., KLEIT, C. J., LASKY, J. J., & CAFFEY, E. M., JR. Combined drug therapy of chronic schizophrenics: A controlled evaluation of placebo, dextroamphetamine, imipramine, isocarboxazid, and trifluoperazine added to maintenance doses of chlorpromazine. *American Journal of Psychiatry*, 1961, 117, 997-1003.
- COOLEY, W. W., & LOHNES, P. R. *Multivariate procedures for the behavioral sciences*. New York: Wiley, 1962.
- ELLSWORTH, R., ARTHUR, G., KROEGER, D., & CHILDERS, B. Comparative validity of staff and relatives' assessment of adjustment for male schizophrenic patients. Paper presented at the meeting of the American Psychological Association, Washington, D. C., September 1967.
- KATZ, M. M., & LYERLY, S. B. Methods of measuring adjustment and social behavior in the community: I. Rationale, description, discriminative validity and scale development. *Psychological Reports*, 1963, 13, 503-535.
- LORR, M. *The Psychotic Reaction Profile manual*. Beverly Hills, Calif.: Western Psychological Services, 1961.
- SMALL, J. G., SMALL, I. F., & GONZALEZ, C. R. The contribution of the informant in psychiatric evaluation. *International Journal of Neuropsychiatry*, 1965, 1, 446-451.
- VESTRE, N. D. Validity data on the Psychotic Reaction Profile. *Journal of Consulting Psychology*, 1966, 30, 84-85.

(Received March 18, 1968)

MATERNAL PERSONALITY CORRELATES OF SONS' CREATIVITY

GEORGE DOMINO¹

Fordham University

California Psychological Inventory scores of 33 mothers of creative high school males, as identified by both teachers' nominations and test scores, were compared with scores of a control group. Mothers of creative Ss exhibit greater self-assurance, initiative, and interpersonal competence; they prefer change and unstructured demands; they are more insightful about others, more tolerant, and value autonomy and independent endeavor. They are, however, less sociable, less conscientious, less dependable, less inhibited, less concerned about creating a favorable impression, and less nurturant and obliging towards others.

Numerous studies have focused on creativity but relatively little is known about what specific parental aspects may be related to filial creativity. The research literature suggests that the home climate of highly creative persons is significantly divergent from that of less creative in several aspects, and it seems reasonable to assume that parental personality characteristics influence a child's creative development.

Weisberg and Springer (1961), in a study comparing psychiatric interview ratings of parents with children's scores on Torrance's tests of creativity, reported that the parents of highly creative children exhibited greater expressiveness with lesser dominance, and were more indifferent towards their children's regressive tendencies. These investigators characterized the emerging family pattern as involving somewhat distant interactions, little clinging to each other for support, little stress on conformity to parental values, mediocre sexual adjustment, paternal exercise of authority, and maternal ambivalence.

MacKinnon (1962, 1965) was able to identify a number of differentiating parental characteristics from the retrospective reports of his creative architects. For example, the parents were seen as exhibiting an extraordi-

nary respect for the child and confidence in his ability to do what was appropriate, while often there was a lack of intense closeness between parent and child.

Dreyer and Wells (1966) studied the instrumental-expressive orientation, role tension, and degree of autonomy-granting of parents of more and less creative children. Parents of more creative children showed greater intrafamilial variance in their rankings of 10 domestic values, and greater role tension reflecting more individual divergence and expression of feeling.

The present report represents one facet of a research project designed to explore parental factors related to filial creativity. This paper will report on maternal personality correlates of sons' creativity. The focus on mothers reflects the assumption that, at least in the American culture, mothers are likely to have more opportunities than fathers to influence their children's psychological growth.

METHOD

Teachers in nine New York State high schools were asked to nominate students who had produced some observable evidence of creativity. These schools are located in primarily middle- to upper-class suburban areas. The specific directions, modified from Schaefer (1967), and reflecting MacKinnon's (1962) definition of creativity, were:

¹ The author is grateful to the Fordham University Computer Center and John Walsh for technical assistance. Anne Anastasi, Marcel Goldschmid, and Harrison Gough gave helpful comments on an earlier draft of this paper. Requests for reprints should be sent to George Domino, Department of Psychology, Fordham University, Bronx, New York 10458.

We are interested in identifying creative students. Please nominate any students you know that you consider as creative. For our purposes, creativity is a process characterized by: 1) originality, novelty, or freshness of approach; 2) adaptiveness to reality in that it must solve a problem, achieve a goal, in general be a reality oriented response;

and 3) the original insight or approach must be developed or elaborated. In order to be nominated then, a student must have actually produced a specific product that can be regarded as both novel (fresh, unusual, imaginative) and of genuine merit or value. The original product might be a drawing, poster, poem, display, composition, invention, experimental design, special project, etc. In identifying creative students, please remember that the novelty and merit of individual products may bear little or no relation to the intelligence of a student, or to whether he is an intellectual, hard-worker, beatnik, or troublemaker. Try not to let any personal feelings influence your evaluation of a student's creativity.

A total of 114 male students were nominated, 37 for artistic activities, 62 for scientific activities, and 15 for both types of activities.

The Remote Associates Test (Mednick & Mednick, 1967) and the Alternate Uses test (Wilson, Christensen, Merrifield, & Guilford, 1960) were administered to those 114 male nominees. To be retained in the creative group, it was arbitrarily decided that a nominee must obtain a raw score of 15 or above on the Remote Associates Test (14 represents the lowest 50th percentile for college students reported by Mednick & Mednick, 1967), and a score of 26 or above on the Alternate Uses test (equivalent to the 54th percentile of college freshmen norms).

Both the teachers' nominations and the stringent test norms were deliberately used in order to identify Ss who could be justifiably called creative. A total of 38 students met the test criteria. Of these 38 Ss, 12 had been nominated for artistic activities, 17 for scientific, and 9 for both. A chi-square analysis of type of activity versus test criteria yielded a non-significant value, indicating that use of the Remote Associates Test and the Alternate Uses test did not discriminate against either the artistic or scientific nominees.

A control group of 38 Ss matched on sex, educational level, grade-point average, and geographical residence, was formed by using school records to identify possible matches, eliminating those nominated as creative or scoring above the cutoff points on the creativity tests. Table 1 gives means and standard deviations of the creativity tests for both control and creative Ss.

Mothers of Ss were then invited to participate in a study of "the influence of the family on children's abilities" by completing the California Psychological Inventory (Gough, 1964a); 33 mothers of creative Ss and 31 mothers of noncreative Ss were tested. These women can be characterized as upper middle class, median age 39, all high school graduates with slightly over one-fourth holding college degrees.

The CPI was used since it has been well validated on large samples of normal Ss and has been used in various studies of creativity (e.g., Helson, 1967; MacKinnon, 1962a, 1962b, 1965). Analyses of mean differences for each of the 18 CPI scales were carried out by means of *t* tests.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF
CREATIVITY TESTS

Ss	Remote Associates Test		Alternate Uses	
	\bar{X}	SD	\bar{X}	SD
Creative	19.2	3.6	29.2	1.9
	12.4	2.7	21.5	2.0

Note.—For Creative, $N = 38$; for Control, $N = 38$.

RESULTS

Twelve mean differences reached statistical significance, 11 at the .01 level and one at the .05 level. The data are presented in Table 2.

Mothers of creative Ss scored significantly higher on Dominance (Do), Capacity for Status (Cs), Tolerance (To), Achievement via Independence (Ai), Psychological-Mindedness (Py), and Flexibility (Fx); they scored significantly lower on Sociability (Sy), Responsibility (Re), Socialization (So), Self-Control (Sc), Good Impression (Gi), and Femininity (Fe).

DISCUSSION

Although the CPI profiles for both groups are above average as compared to the general population, the results clearly indicate that mothers of creative Ss differ significantly from mothers of noncreative Ss on several dimensions of personality. Mothers of creative Ss exhibit greater self-assurance and initiative (Do) and interpersonal competence (Cs); they prefer change and unstructured demands (Fx); they are more insightful about others (Py), but at the same time more tolerant (To), and they value autonomy and independent endeavor (Ai). They are, however, less sociable (Sy), exhibit less responsibility (Re) and social probity (So), are less inhibited (Sc), less concerned about creating a favorable impression (Gi), and less nurturant and obliging toward others (Fe).

The general picture generated by this CPI mean profile is that of a highly independent and capable woman perceptive in her interpersonal relationships but little disposed to worry about the impression she creates. She

TABLE 2
CPI SCORES OF MOTHERS OF CREATIVE AND NONCREATIVE ADOLESCENTS

CPI Scale	Creative		Noncreative		<i>t</i>
	\bar{X}	<i>SD</i>	\bar{X}	<i>SD</i>	
Dominance (Do)	30.2	5.2	26.1	4.9	3.25**
Capacity for Status (Cs)	23.8	3.1	19.2	3.3	5.74**
Sociability (Sy)	22.7	4.7	26.9	4.6	3.61**
Social Presence (Sp)	35.4	5.3	35.2	5.2	0.07
Self-Acceptance (Sa)	22.9	3.7	22.6	2.9	0.36
Well-Being (Wb)	38.8	3.6	39.5	4.4	0.69
Responsibility (Re)	29.1	4.5	34.6	3.9	5.48**
Socialization (So)	35.6	5.1	42.1	4.6	5.36**
Self-Control (Sc)	27.5	7.2	35.5	6.1	4.81**
Tolerance (To)	27.3	4.3	24.1	3.9	3.12**
Good Impression (Gi)	16.6	5.5	20.3	6.1	2.54*
Communality (Cm)	24.5	1.8	25.1	1.3	1.53
Achievement via Conformance (Ac)	29.3	4.2	29.4	3.7	0.10
Achievement via Independence (Ai)	24.3	3.8	21.1	2.9	3.80**
Intellectual Efficiency (Ie)	44.8	5.2	42.9	4.1	1.62
Psychological-Mindedness (Py)	15.8	2.9	12.6	2.4	4.82**
Flexibility (Fx)	14.9	3.6	11.3	3.3	4.17**
Femininity (Fe)	22.4	3.5	24.8	3.7	2.66**

Note.—For Creative; *N* = 33; for Noncreative, *N* = 31.

**p* < .05.

***p* < .01.

can adapt to change and variety, and in fact is restive when faced with rules and constraints. She is free in her expression of impulse, but perhaps not always dependable.

These themes emerging from the CPI profile have in many cases been previously reported as characteristic of creative persons (cf. Barron, 1965; Gough, 1964b; MacKinnon, 1963).

One theme relates to interpersonal endeavor. These women have the assurance, the versatility and ascendancy necessary to achieve a position of high community visibility. Yet they do not. Of the 33 mothers of creative Ss, only four were actively involved in some community activity, as opposed to 17 of the 31 mothers of noncreative Ss ($\chi^2 = 15.17$, $p < .01$). They are also moderately indifferent to social demands: of the six CPI scales addressed to various aspects of socialization and control, they score significantly lower on four (Re, So, Sc, Gi). These results are consonant with those of Dreyer and Wells (1966) who found that the mothers of creative children were much less concerned with a place in the community and more concerned with everyday interests. MacKin-

non (1961) reported that his creative architects revealed less desire to be included in group activities than any other group studied by him and his coworkers. Yet, when these architects chose to participate they did so with consummate skill.

A second theme concerns interpersonal closeness. These mothers are perceptive individuals, alert to the needs of others, informal, tolerant, and observant (To, Py, Fx). Yet, these socially desirable characteristics are counterweighted by a degree of indifference and detachment (Gi, Sy). These results are consonant with the recollections of MacKinnon's architects who reported a lack of intense closeness with their parents. They also agree with the findings of Getzels and Jackson (1962) that mothers of more creative children were less observant of failings in their children than were mothers of less creative.

A third theme reflects impulse expression, as shown primarily by the lower mean on the Sc scale, but also supported by lower means on So and Re. Barron (1957) has seen this factor as indicating the potentiality of the pleasure principle occasionally to become ascendant over the reality principle. One might

speculate that this reduction of self-control allows the person to regress in the service of the ego, and that mothers of creative children are themselves more creative than their peers.

A fourth theme concerns initiative. These women exhibit certain personality characteristics which in our culture tend to be labeled as more masculine. This can be seen in the significantly higher means on the Do and Ai scales, as well as by the lower mean on the Fe scale. Helson (1967) has reported that a masculine component is an important factor in the syndrome underlying creativity in women.

A final theme has to do with independence. Independence in thought and action as a characteristic of creative persons has been reported by several investigators. Dreyer and Wells (1966) reported that the mothers of high creative children stressed emotional security, leading to greater confidence and independence in their offspring. MacKinnon (1961) found that creativity in architects correlated highly with self-assertive independence. The present CPI findings are consonant in revealing these mothers to be more independent, autonomous, and self-reliant.

It is interesting to note a possible (but superficial) discordance in these themes: high social poise with little affiliation; interpersonal insight with indifference; achievement drive and masculine initiative with biological femininity. Thus there is a balance between contrasting dynamic forces, in which positive and negative characteristics are both allowed expression. Jung's concept of the transcendent function which accomplishes the successful integration of opposites is relevant here.

The findings presented here indicate quite clearly that mothers of creative adolescents possess personality characteristics not only significantly different from those of mothers of noncreative adolescents, but quite similar to those observed among creative adults. Although there is no direct evidence, it is reasonable to postulate that the mothers of creative Ss are themselves more creative than

the general population, and that both their personality characteristics and their creativity are evocative of greater creativity in their children.

REFERENCES

- BARRON, F. Originality in relation to personality and intellect. *Journal of Personality*, 1957, 25, 730-742.
- BARRON, F. The psychology of creativity. In, *New directions in psychology: II*. New York: Holt, Rinehart & Winston, 1965.
- DREYER, A., & WELLS, M. Parental values, parental control, and creativity in young children. *Journal of Marriage and the Family*, 1966, 28, 83-88.
- GETZELS, J., & JACKSON, P. *Creativity and intelligence*. New York: Wiley, 1962.
- GOUGH, H. G. *California Psychological Inventory manual*. (Rev. ed.) Palo Alto, Calif.: Consulting Psychologists Press, 1964. (a)
- GOUGH, H. G. Identifying the creative man. *Journal of Value Engineering*, 1964, 2, 5-12. (b)
- HELSON, R. Personality characteristics and developmental history of creative college women. *Genetic Psychology Monographs*, 1967, 76, 205-256.
- MACKINNON, D. W. The personality correlates of creativity: A study of American architects. In Gerhard S. Nielsen (Ed.), *Proceedings of the XIV International Congress of Applied Psychology, Copenhagen, 1961*. Vol. 2. Copenhagen: Munksgaard, 1962. (a)
- MACKINNON, D. W. The nature and nurture of creative talent. *American Psychologist*, 1962, 17, 484-495. (b)
- MACKINNON, D. W. Creativity and images of the self. In R. W. White (Ed.), *The study of lives*. New York: Atherton, 1963.
- MACKINNON, D. W. Personality and the realization of creative potential. *American Psychologist*, 1965, 20, 273-281.
- MEDNICK, S. A., & MEDNICK, M. T. *Examiner's manual for the Remote Associates Test*. Boston: Houghton Mifflin, 1967.
- SCHAEFER, C. E. Biographical inventory correlates of scientific and artistic creativity in adolescents. Unpublished doctoral dissertation, Fordham University, 1967.
- WEISBERG, P. S., & SPRINGER, K. J. Environmental factors in creative function. *Archives of General Psychiatry*, 1961, 5, 554-564.
- WILSON, R. C., CHRISTENSEN, P. R., MERRIFIELD, P. R., & GUILFORD, J. P. *Alternate uses: Manual*. Beverly Hills, Calif.: Sheridan Supply, 1960.

(Received March 28, 1968)

IDENTIFYING NEUROLOGICALLY IMPAIRED CHILDREN THROUGH A TEST OF AUDITORY PERCEPTION

DAVID A. SABATINO¹

The Catholic University of America School of Education

The purpose of this study was to determine if an experimental Test of Auditory Perception (TAP) could discriminate between neurologically impaired and normal children. The relationships were established between each subtest of the Test of Auditory Perception and the Bender Visual Motor Gestalt Test (BVMGT). When the TAP was administered under normal testing procedures, four of the six subtests discriminated the experimental from the control Ss at the .01 level. All of the TAP subtests discriminated the neurologically impaired from normal children at the .01 level under conditions of background noise. The correlations between the TAP subtests were positive but low in value. Negative correlations were found between each of the TAP subtests and the BVMGT.

The behavioral sciences have developed many tests of visual perception as means of identifying children who are unable to interpret environmental stimuli as meaningful information. Herbert (1964), in reviewing the psychological literature, found the Bender Visual Motor Gestalt Test (BVMGT) to be the most popular test of visual motor perception. Sundburg (1961), found the BVMGT to be the fifth most commonly used clinical test.

Bender (1938), assumes that visual perception based on Gestalt principles can: (a) differentiate development levels, and (b) differentiate neurologically impaired from non-neurologically impaired Ss. It seems in reviewing the literature (Barnes, 1950; Benton, 1962; Clements & Peters, 1962) that the behavioral sciences have established clinical evidence to support a relationship between visual-motor perception and minimal neurological impairment.

In contrast, research efforts attempting to study the auditory perceptual function of children suspected of neurological impairments are limited. To quote Gallagher (1960), "in contrast to the visual-motor areas, other areas of sensory perception have been almost neglected [p. 19]."

Strong (1964) factored 92 test and subtest variables utilizing computer-based factor analysis. He found that the auditory decoding and auditory-vocal automatic subtests of the

Illinois Test of Psycholinguistic Abilities (ITPA; McCarthy & Kirk, 1961) correlated .58 and .51, respectively, with the "brain damage syndrome." He found these two tests to correlate even lower with academic achievement. More importantly, when factored for auditory memory, these two subtests correlated the lowest of eight other tests of this skill. The reliability coefficients were in the low forties. None of the other seven subtests of the ITPA correlated well with any criteria.

Hasterok (1964) found that visual and auditory difficulties, as identified in ITPA profiles, were unsuited to answering the important question of the relationship between learning problems and perceptual or sensory problems. He concluded that children cannot be matched for comparison purposes on the basis of sensory or perceptual problems. The problem is that the ITPA was designed as a measure of psycholinguistic skills and not perceptual functions.

PURPOSE

The problem within the initial design of this study became that of developing an experimental test of auditory perception, since comprehensive measures of this function were not available.

The purpose of this study was to determine the ability of the BVMGT and Experimental Test of Auditory Perception (TAP) to differentiate children with neurological impairment from children with no known neurological impairment. In other words, the primary

¹Requests for reprints should be sent to David A. Sabatino, School of Education, The Catholic University of America, Washington, D. C. 20017.

TABLE 1
MATCHING CRITERIA FOR THE EXPERIMENTAL AND CONTROL SUBJECTS

Group	Mean chronological age in years and months		Mean years in school	Sex		Verbal IQ	
	M	Range		Boys	Girls	M	Range
Experimental	8.3*	6 to 12-5	3.4	19	11	105.5**	87-130
Control	8.9*	6 to 12-4	3.7	18	12	106.9**	88-133

* $p < .05$.

** $p < .01$.

dimension of this research was to obtain data concerning the means by which a commonly used test of visual-motor perception and an experimentally constructed TAP would identify children previously known to be neurologically impaired (experimental Ss) from those with no such impairment (control Ss).

METHOD

Thirty Ss with known brain damage, minimal in extent, and 30 Ss with no known cerebral pathology were used as experimental and control Ss. The Ss ranged in age from 6 years 1 month to 12 years 7 months. The brain-damaged children were selected on the basis of pediatric-neurological examination, including electroencephalograms, from a large population of children referred to Children's Hospital Learning Disability Center, Columbus, Ohio.

The children selected as Ss were free from physical, health, gross motor, or social-personal problems. All the children were given sensory hearing screening tests of 0 decibels (ISO) on pure-tone air conduction tests at 500, 1,000 and 2,000 Hz.

The primary purpose was to determine if the experimentally constructed TAP and the BVMGT would discriminate between these two groups. The experimental and control groups were matched on age, sex, and verbal IQ. Table 1 shows the matching criteria and accompanying level of significance.

Construction of the Experimental Test of Auditory Perception

The TAP was recorded on electromagnetic tape by placing the tape recorder (Wollensak-T-1500) into a sound-treated chamber (Industrial Acoustics Company, Model 402). The stimulus output was established at 60 decibels in order that the inter-S test consistencies could be maintained. The original test tape recording was made with a known output, free from any interfering noise in the background. A second TAP containing a background of broad spectra sawtooth noise was constructed by recording

the original test tape and controlled noise simultaneously. The signal-to-noise ratio was 60/54 decibels.

The content of the TAP was selected to be the simplest and purest measures of auditory perception for each hypothetical stage. The difficulties in assembling such a test were:

1. The perceptual stimuli must be received solely by the auditory mechanism and not by other sensory and perceptual systems or the function measured would be an integrative one.

2. The response mechanism must not require a complex response or the perceptual measurement would become one of S's motor or vocal response.

The following are the subtests of the TAP:

1. *Auditory Recognition* is the meaningful recognition of sounds and words. The subtests used to determine auditory recognition were: (a) *A Speech Sound Discrimination Test* of paired nonsense syllables which were identified as alike or different. (b) *A Word Discrimination Test* which required S to distinguish a meaningful word presented aurally from five nonsense words which sounded similarly.

2. *Auditory Retention* is the recall of connected units of language. Speech sounds are first recognized as meaningful, then stored in sequential order and retained if further association of their meaning is to take place. (a) *The Immediate Recall of Rote Digital Units*, administered at 1-second intervals requiring S to point to graphic number representations at the conclusion of the sequence. (b) *The Memory for Connected Speech*, was based on a sentence memory test which required an immediate meaningful recall of connected words.

3. *Auditory Integration* is the ability of the listener to bring together disassociated auditory stimuli. The integrative function of language has been shown to be approximated by having children reproduce rhythmic patterns. (a) *Rhythmic Structures*. The child was requested to duplicate tapping responses to pre-recorded patterns.

4. *Auditory Comprehension* is the summative ability of the child to receive, retain, and integrate units of aural language. It is the final, complex auditory perceptual act based on all the functions previously established. (a) *The Test of Auditory Comprehension* required the child to listen to a story and re-

TABLE 2

DISCRIMINATION OF SUBJECTS BY THE BVMGT AND THE TAP AT THE STATISTICAL CRITERION UNDER NORMAL TESTING CONDITIONS AND CONDITIONS OF BACKGROUND NOISE

Test	Background noise		Normal testing conditioning	
	Experimental	Control	Experimental	Control
BVMGT	8	2	8	2
TAP	29	4	23	5

spond to questions which generally called for factual one-word responses of yes or no.

Experimental Assessment and Scoring Procedures

The two tests of auditory perception were administered on different days because it was necessary to randomize the administration of the tests among Ss. Both tests (the original test recording and the one with noise in the background) had the same content.

Three of the subtests merely required yes, no, or pointing responses. One subtest required a tapping response, and two required oral responses. These were the subtests of word recognition and comprehension. Both of these subtests confined the length of response to one word. The dichotomous response choices with limited oral interaction reduced the possibility of any scoring problem due to an articulation problem in a child's speech.

The BVMGT was administered by placing a test card in the upper-left-hand corner of an $8\frac{1}{2} \times 11$ -inch sheet of plain white paper. The card and paper were both held in position in such a manner that the examiner had full control of the test material and could maintain a consistency between administration procedures for each S. The BVMGT was scored using the Koppitz (1964) developmental scoring system.

RESULTS

A statistical criterion was established at the first minus standard deviation ($-1 SD$) to differentiate the experimental and control S responses to the BVMGT and the TAP. The adoption of this statistical criterion permits Ss to be categorized in regard to their number of errors on the BVMGT and TAP. Table 2 shows the number of experimental and control Ss that were functioning more poorly (more errors) than the statistical criterion would permit for both TAP and the BVMGT. The data indicate that the BVMGT identified eight of the experimental Ss, while the TAP with background noise identified 29 Ss, and 23 under the original testing conditions.

The relationship between each TAP subtest and the BVMGT was determined by multiple correlational analyses. All of the relationships were negative. Table 3 shows the correlational values for each S group on each subtest. The degree of negative correlation runs from $-.19$ to $-.64$. This means that none of the TAP subtests assess the same behavior as the BVMGT.

It is apparent from the data involved in Tables 2 and 3 that the human functions assessed by the BVMGT are not the same as assessed by the TAP in either normal or neurologically impaired children. In fact, the data suggests that the routine assessment of visual perception may contribute to our overlooking a very large group of children with serious auditory perceptual impairment. The data establishes that each of the TAP subtests assessed a different function than does the BVMGT. Therefore, it became important

TABLE 3

CORRELATIONAL VALUES BETWEEN THE BVMGT AND EACH TAP SUBTEST

Test correlations	Covariance		Correlation	
	Experimental	Control	Experimental	Control
BVMGT \times Auditory Recognition (sounds)	-.7332	-.1903	-.562	-.368
BVMGT \times Auditory Recognition (words)	-.1163	-.1078	-.348	-.397
BVMGT \times Auditory Retention (digits)	-.3713	-.2731	-.384	-.37
BVMGT \times Auditory Retention (sentences)	-.1011	-.1213	-.192	-.368
BVMGT \times Auditory Integration (integration)	-.1904	-.4508	-.151	-.492
BVMGT \times Auditory Comprehension	-.1871	-.5772	-.644	-.244

TABLE 4

LEVELS OF SIGNIFICANCE BETWEEN THE MEANS OF THE EXPERIMENTAL AND CONTROL GROUPS FOR EACH TAP SUBTEST ADMINISTERED UNDER CONDITIONS OF BACKGROUND NOISE

Subtests of Auditory Perception	<i>t</i>
Auditory Recognition (sounds)	8.017*
Auditory Recognition (words)	9.52565*
Auditory Retention (digits)	6.396*
Auditory Retention (sentences)	7.235*
Auditory Integration (tapping)	3.790*
Auditory Comprehension (stories)	4.901*

* $p < .01$.

to determine which subtests of the TAP could most aptly discriminate the experimental and control Ss under test administrations of background-noise and original testing (no-background-noise) conditions.

The *t*-test values indicated that under conditions of background noise, each of the TAP subtests discriminated between the experimental and control Ss at the .01 level as significant. The results are shown in Table 4.

Under the original test conditions (no-background-noise) the subtests of auditory retention (digits) and auditory integration (tapping responses) were unable to discriminate between the experimental and control Ss at a significant level. All the other subtests revealed significant differences at the .01 level. This means that nonsense words or sounds, which have no direct meaning for language, in sequence or interpretation, are a poor means of assessing auditory perception. This probably is in keeping with the standard definition of perception, which posits the interpretation of incoming sensory stimuli in order that it may be correctly relayed to higher cortical centers. Repeating digits in rote fashion and listening to tapping patterns are not tasks that require interpretation of an auditory stimulus.

CONCLUSION

Clinically, the BVMGT is very commonly used to determine if possible neurological damage is present in children. The purpose of this study was to determine if it would discriminate children with known neurological impairment from children with no known neu-

TABLE 5

LEVELS OF SIGNIFICANCE BETWEEN THE MEANS OF THE EXPERIMENTAL AND CONTROL GROUPS FOR EACH TAP SUBTEST ADMINISTERED UNDER NORMAL TESTING CONDITIONS

Subtests of Auditory Perception	<i>t</i>
Auditory Recognition (sounds)	4.068*
Auditory Recognition (words)	3.5898*
Auditory Retention (digits)	1.9447
Auditory Retention (sentences)	3.3374*
Auditory Integration (tapping)	1.3951
Auditory Comprehension (stories)	4.5938*

* $p < .01$.

rological impairment better than an experimental TAP.

The results indicate that the experimental TAP discriminated the neurologically impaired from normal Ss at a significantly higher level than did the BVMGT. Furthermore, the TAP correlates negatively with the BVMGT. This indicates that the TAP and BVMGT do not predict the same perceptual behaviors.

The TAP had much more discrimination power when administered in background noise. This suggests that perceptually disturbed children have difficulty separating ambient auditory background (noise) from a meaningful stimulus signal. It is interesting to note that only the TAP subtests that contained meaningful language were affected by ambient background noise.

REFERENCES

- BARNES, T. C. EEG validity of Rorschach, Hunt, and Bender Gestalt Test. *American Psychologist*, 1950, 5, 332.
- BENDER, L. A Visual-Motor Gestalt Test and its clinical use. *The American Orthopsychiatric Association Research Monograph*, 1938, No. 2.
- BENTON, A. L. Behavioral indices of brain injury in school children. *Child Development*, 1962, 33, 199-208.
- CLEMENTS, S. D., & PETERS, J. E. Minimal brain dysfunction: In the school age child: Diagnosis and treatment. *Archives of General Psychiatry*, 1962, 6, 185-197.
- GALLAGHER, J. J. *The tutoring of brain injured mentally retarded children*. Springfield, Ill.: Charles C Thomas, 1960.
- HASTEROK, G. S. The learning performance of retarded children with visual and auditory sense modality disabilities. Unpublished doctoral dissertation, University of Illinois, 1964.

- HERBERT, M. The concept and testing of brain damage in children: A review. *Journal of Child Psychology and Psychiatry*, 1964, 5, 197-216.
- KOPPITZ, E. *The Bender Gestalt Test for young children*. New York: Grune & Stratton, 1964.
- MCCARTHY, J. J., & KIRK, S. A. *Illinois Test of Psycholinguistic Abilities: Experimental edition*. Urbana, Ill.: Institute for Research on Exceptional Children, University of Illinois, 1961.
- STRONG, R. The identification of primary school age children with learning handicaps associated with minimal brain disorders. Unpublished doctoral dissertation, University of Utah, 1964.
- SUNDBURG, N. The practice of psychological testing in clinical service in the United States. *American Psychologist*, 1961, 16, 83.

(Received March 25, 1968)

EFFECTS OF RATER LEVEL OF FUNCTIONING AND EXPERIENCE UPON THE DISCRIMINATION OF FACILITATIVE CONDITIONS

JOHN R. CANNON¹ AND ROBERT R. CARKHUFF²

State University of New York at Buffalo

The ratings of 80 Ss, at four different levels of experience, at each of two levels of facilitative interpersonal functioning (moderately high and low), were analyzed. Both level of functioning and experience were found to have significant effects on accuracy of ratings, with level of functioning a more critical index of ability to discriminate levels of facilitative conditions.

With the extended use of scales for rating interpersonal processes, concern with psychometric rigor has led investigators of rating methodology to focus almost exclusively upon the reliability of ratings while neglecting the crucial questions of accuracy or validity (Auld & Murray, 1955; Shoben, 1956; Taft, 1959, 1966; Vargas, 1954).

For example, the work of Shapiro (1968) indicates that "untrained raters are able to differentiate high and low levels of psychotherapeutic behavior in a manner which is similar to that of trained raters [p. 88]." However, this research does not indicate whether or not the ratings of either group of Ss are related to any validity criteria. Further, while suggesting, in effect, that everyone can discriminate gross levels of functioning, Shapiro's (1968) study leaves the reader without cues as to which raters can make the most accurate or valid predictions.

A review of the literature dealing with raters and their personal characteristics reveals that two variables have been of primary concern to investigators of general, person-perception rating methodology: (a) the clinical experience or training of the rater, and (b) the rater's adjustment or level of effective functioning. While the results of investigations of the effects of experience are, at best, equivocal (e.g., Amble & Moore, 1966;

Arnhoff, 1954; Greenwood & McNamara, 1967; Jones, 1957; Oskamp, 1962; Watley, 1967), there is substantial evidence that rater adjustment or level of functioning has significant effects on the accuracy of person-perception ratings (Reid & Snyder, 1947; Taft, 1955; Vingoe & Antonoff, 1968; Watley, 1967).

The present study investigates the effects of both rater level of facilitative interpersonal functioning and rater level of experience upon the accuracy of discriminations of the facilitative conditions of empathy, respect, genuineness, concreteness, self-disclosure, confrontation, and immediacy (Berenson & Mitchell, 1968; Carkhuff & Berenson, 1967; Rogers, 1967; Truax & Carkhuff, 1967).

METHOD

Subjects

Eighty Ss were selected from a pool of 251 Ss whose written responses to standard client stimulus expressions had been assessed by gross rating scales of facilitative interpersonal communication (see Figure 1). The final 80 Ss included 20 Ss at each of four experience levels: (a) experienced counselors and psychotherapists; (b) graduate trainees in counseling and psychotherapy; (c) undergraduates with experience in a helping role; and (d) undergraduates with no experience in a helping role. The 20 Ss at each experience level included the 10 highest rated and the 10 lowest rated respondents at that level of experience.

Procedure

Communication. A tape comprised of 16 client statements of no more than 2 minutes duration each (total playing time of approximately one-half hour) was prepared to present as stimulus material to elicit helper responses from Ss (Carkhuff, in press). The client stimulus expressions tape, representing

¹ Now at the Division of Counseling, Pennsylvania State University. Requests for reprints should be sent to John R. Cannon, 504 Beaumont Drive, State College, Pennsylvania.

² The authors wish to thank Tom Collingwood and Lauri Renz for their assistance in organizing and analyzing the data and Ted Friel and Daniel Kratochvil for their assistance in rating.

The facilitator is a person who is living effectively himself and who discloses himself in a genuine and constructive fashion in response to others. He communicates an accurate empathic understanding and a respect for all of the feelings of other persons and guides discussions with those persons into specific feelings and experiences. He communicates confidence in what he is doing and is spontaneous and intense. In addition, while he is open and flexible in his relationships with others, in his commitment to the welfare of the other person he is quite capable of active, assertive and even confronting behavior when it is appropriate.

You will hear a number of excerpts taken from therapy sessions. Rate each therapist response to each excerpt 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, or 5.0 using the continuum shown below.

1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
None of these conditions are communicated to any noticeable degree in the person.		Some of the conditions are communicated and some are not.		All conditions are communicated at a minimally facilitative level.		All of the conditions are communicated, and some are communicated fully.		All are communicated fully, simultaneously and continually

FIG. 1. Gross ratings of facilitative interpersonal functioning.

client statements from live counseling, crossed three dominant affect areas with five dominant content areas. The affect areas included the following: (a) depression-distress; (b) anger-hostility; (c) elation-excitement. The content areas included the following: (a) social-interpersonal; (b) educational-vocational; (c) child-rearing; (d) sexual-marital; (e) confrontation of counselor. The excerpts were arranged in such a way that each affect area was matched with each content area. In addition, a silence was included as a stimulus for eliciting Ss' responses.

All 251 Ss were presented with the 16 client statements and asked to write out their own best responses to each of the client expressions of feeling and content. Communication ratings of Ss' written responses were obtained from two experienced raters who had established high intra- and interrater reliabilities (ranging from .80 to .99) in previous research employing individual scales for the counselor dimensions (empathy, respect, genuineness, concreteness, self-disclosure, confrontation, and immediacy). In the present research project, these raters were asked to employ the gross ratings form for assessing inter-

personal functioning (see Figure 1), the same form that Ss were to employ in making their discriminations. The rate-rater reliabilities of the two experienced raters employing the gross ratings form were .95 and .93, and the interrater reliability was .89.

Taking into consideration only Ss' own mean levels of communication (as rated by the two experienced raters), the 10 highest and 10 lowest rated Ss at each level of experience were selected for the final S groups.

Discrimination. A second tape, comprised of the 16 client stimulus expressions and four possible counselor responses to each expression, was prepared to present for Ss' ratings of counselor responses. The counselor responses, representing counselor statements from live counseling, were designed so that the four responses to each client expression demonstrated varying levels of facilitative interpersonal communication (Carkhuff, in press).

The Ss' discrimination scores were established by determining the mean absolute deviation of an S's ratings from the consensus ratings of two experts (different from the raters of Ss' communication) who, in previous outcome studies, had demonstrated a

TABLE 1

MEANS AND STANDARD DEVIATIONS OF COMMUNICATION AND DISCRIMINATION SCORES

Group	Ss	Communication		Discrimination	
		Ratings on 5-point scales		Deviations from experts	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
I	High-level experienced counselors and therapists	2.79	.45	.45	.15
II	Low-level experienced counselors and therapists	1.44	.28	.78	.20
III	High-level graduate trainees	2.93	.23	.49	.17
IV	Low-level graduate trainees	1.58	.09	1.03	.23
V	High-level experienced undergraduates	2.25	.20	.94	.20
VI	Low-level experienced undergraduates	1.18	.06	1.14	.17
VII	High-level inexperienced undergraduates	2.13	.08	.94	.16
VIII	Low-level inexperienced undergraduates	1.17	.05	1.19	.08
	All low-level Ss	1.34	.23	1.04	.24
	All high-level Ss	2.52	.44	.71	.29
	All Ss	1.93	.69	.87	.31

Note.—*N* = 10 for each group; totals, all low-level Ss, *N* = 40, all high level Ss, *N* = 40, all Ss, *N* = 80.

high degree of predictive validity in their use of the rating scales (Carkhuff, 1968, 1969, in press; Carkhuff & Berenson, 1967).

RESULTS

Group means for both communication (Ss' levels of facilitative interpersonal functioning) and discrimination (deviations from criterion ratings) appear in Table 1. It can be seen that at each experience level, the low-level group was functioning approximately one to one and one-half communication levels lower than the high-level group.

The 4 × 2 analysis of variance (see Table 2) indicated that Ss' level of functioning, Ss' experience, and the interaction between these two variables all were significant sources of variance in discrimination scores.

Simple effects of level of functioning were evaluated with *t* tests, and it was found that at every level of experience, the high-level functioning Ss were significantly ($p < .05$) more accurate in their ratings than the low-level Ss.

In order to control for the confounding of experience with level of functioning, an analysis of covariance was conducted to assess the independent effect of experience on discrimination (covaried for level of functioning). The *F* computed from this analysis was highly

significant ($p < .001$), and, therefore, separate analyses of covariance and *t* tests based on the adjusted means and adjusted error terms were conducted at each of the two levels of functioning.

The results of these analyses revealed the following: (a) At both high and low levels of functioning, experience had a significant ($p < .01$) effect on accuracy of ratings; (b) among high-level functioning Ss, there were significant ($p < .05$) differences in rating accuracy between more experienced Ss (Groups I and III) and undergraduate Ss (Groups V and VII), but no significant differences either between the experienced counselors and the graduate students (I and III)

TABLE 2

SUMMARY OF OVERALL ANALYSIS OF VARIANCE OF DISCRIMINATION SCORES

Source	<i>df</i>	<i>MS</i>	<i>F</i>
Experience	3	.9492	27.63**
Level of functioning (communication)	1	2.2005	64.05**
Experience × Level	3	.1168	3.40*
Within cells	72	.0344	
Total	79		

* $p < .05$.

** $p < .001$.

TABLE 3

DISCRIMINATION DIFFERENCES BETWEEN HIGH-LEVEL FUNCTIONING SUBJECTS AND LOW-LEVEL FUNCTIONING SUBJECTS (VALUES OF t)

Group	II Low-level experienced counselors	IV Low-level graduate trainees	VI Low-level experienced undergraduates	VIII Low-level inexperienced undergraduates
I High-level experienced counselors	-3.98***	-7.00***	-8.32***	-8.93***
III High-level graduate trainees	-3.50***	-6.51***	-7.84***	-8.44***
V High-level experienced undergraduates	+1.93	-1.09	-2.41*	-3.02**
VII High-level inexperienced undergraduates	+1.93	-1.09	-2.41*	-3.02**

Note.—Negative numbers indicate high-level group more accurate than low-level group. Positive numbers indicate low-level group more accurate than high-level group.

* $p < .05$.
 ** $p < .01$.
 *** $p < .001$.

or between the experienced and inexperienced undergraduates (V and VII); (c) among low-level functioning Ss; experienced counselors (Group II) were significantly ($p < .05$) more accurate in their ratings than all other groups, with no significant differences in rating accuracy between graduate students (IV), experienced undergraduates (VI), and inexperienced undergraduates (VIII).

Considering the results of the simple-effects tests for both experience and level of functioning, it appears that the significant effect of the interaction of Experience \times Level of Functioning (Table 2) reflects the differential effects of experience at the two levels of functioning rather than differential effects of level of functioning at the four levels of experience.

Such an interpretation provides tentative support for the assumption that level of functioning is a more critical indicator of rating accuracy than level of experience; and, in order to further investigate the relative effects of experience and level of functioning, each group of high-level functioning Ss was compared to every group of low-level functioning Ss. Table 3 presents the numerical values and probabilities of the t 's computed from these comparisons. Perusal of the entries in Table 3 reveals the following: (a) two groups of high-level functioning Ss (I and III) were significantly ($p < .001$) more accurate in their ratings than every group of low-level functioning Ss; (b) each of the remaining two groups of high-level functioning Ss (V and VII) was significantly ($p < .05$) more accurate than two of the groups of low-level

functioning Ss (VI and VIII), and not significantly different from the other two low-level groups (II and IV); (c) none of the groups of low-level functioning Ss, regardless of experience, was significantly more accurate in rating than any of the groups of high-level functioning Ss.

Finally, correlations were computed between level of functioning (communication) and discrimination at both high and low levels of functioning with the following results: (a) The Pearson r between communication and discrimination for all low-level functioning Ss was $-.26$ (ns); (b) the r between communication and discrimination for all high-level functioning Ss was $-.76$ ($p < .001$). Due to the methods of scoring these variables, a negative correlation indicates that the higher an S's level of functioning, the more accurate are his ratings.

DISCUSSION

The present findings indicate that rater level of interpersonal functioning and rater level of experience have significant effects upon the accuracy of ratings of interpersonal processes, with rater level of functioning appearing to be the more critical indicator of rating accuracy. At this point in the development of process research, the results lend support to proponents of the criterion of meaning who suggest that the validity of rating scales is largely a function of the particular raters who employ them (Carkhuff, 1968; Lehmann, Ban, & Donald, 1965; Marsden, 1965).

The results appear to be inconsistent with those of Shapiro (1968) and others who indicate that neophytes rate as effectively as trained raters, and do so with gross rating scales that involve minimal effects of experience. However, the present findings may account for the results of such previous studies with the possibility that both the neophytes and the experienced raters may have been functioning at the same levels, whether high, moderate, or low. For example, in the present study, high-level experienced undergraduates were very similar in discrimination to high-level inexperienced undergraduates, and low-level experienced undergraduates were very similar in discrimination to low-level inexperienced undergraduates.

On the other hand, the results of the present research also imply that accuracy of discriminations may be largely determined by level of functioning of the rater regardless of whether he is a therapist, trainee, client, or undergraduate student; although, to be sure, one may be able to make diminishing probability statements in this regard. For example, an experienced therapist in the low group (II) may discriminate neither as well as a graduate trainee (III) nor significantly better than experienced and inexperienced undergraduates (V and VII).

The representativeness of the present samples is supported by the essential replication of previous findings of average levels of interpersonal functioning at different experience levels (Carkhuff, 1969; Carkhuff & Berenson, 1967). In addition, the differential findings concerning the intercorrelations of communication and discrimination within groups of high-level and low-level functioning Ss are consistent with those of Carkhuff, Kratochvil, and Friel (1968) indicating the independence of communication and discrimination among low-level functioning Ss and a high degree of relationship between these variables among high-level functioning Ss. The situation, it appears, is not unlike the insight-action discrepancy often found with clients in counseling. That is, in general, the evidence indicates that nothing relates very much to anything else among low-level functioning Ss, while there are very consistent and stable relationships between relevant variables

among high-level functioning Ss (Anthony, 1968; Carkhuff, 1969, in press; Greenbaum, 1968; Greenberg, 1968).

In summary, there is now empirical as well as experiential evidence to indicate that prospective raters of interpersonal processes should be selected on the basis of two criteria: (a) demonstrated high levels of functioning in the relevant area; and (b) experience in the relevant area. Of these two selection indexes, the evidence suggests that level of functioning is the more critical.

REFERENCES

- AMBLE, B. R., & MOORE, R. The influence of a set on the evaluation of psychotherapy. *American Journal of Orthopsychiatry*, 1966, 36, 50-56.
- ANTHONY, W. A. The effects of rehabilitation counselor training upon trainee communication, discrimination and helping attitudes. Unpublished doctoral dissertation, State University of New York at Buffalo, 1968.
- ARNHOFF, F. N. Some factors influencing the unreliability of clinical judgments. *Journal of Clinical Psychology*, 1954, 10, 272-275.
- AULD, F., JR., & MURRAY, E. J. Content-analysis studies of psychotherapy. *Psychological Bulletin*, 1955, 52, 377-395.
- BERENSON, B. G., & MITCHELL, K. Confrontation in counseling and life. Unpublished manuscript, State University of New York at Buffalo, 1968.
- CARKHUFF, R. R. Counseling and human relations: A primer for lay and professional helpers. State University of New York at Buffalo, 1968. (Mimeo)
- CARKHUFF, R. R. Helper communication as a function of helpee affect and content. *Journal of Counseling Psychology*, 1969, 16, 126-131.
- CARKHUFF, R. R. The prediction of the effects of teacher-counselor training: The development of communication and discrimination selection indexes. *Counselor Education and Supervision*, in press.
- CARKHUFF, R. R., & BERENSON, B. G. *Beyond counseling and therapy*. New York: Holt, Rinehart & Winston, 1967.
- CARKHUFF, R. R., KRATOCHVIL, D., & FRIEL, T. Effects of professional training: Communication and discrimination of facilitative conditions. *Journal of Counseling Psychology*, 1968, 15, 68-74.
- GREENBAUM, S. Response disposition to counselor responses of low-level functioning communicators. Unpublished manuscript, State University of New York at Buffalo, 1968.
- GREENBERG, B. Differential effects of responding verbally and in writing to client stimulus expressions and in a standard client interview situation. Unpublished manuscript, State University of New York at Buffalo, 1968.
- GREENWOOD, J. M., & McNAMARA, W. J. Interrater reliability in situational tests. *Journal of Applied Psychology*, 1967, 51, 101-106.

- JONES, N. F. Context effects in judgment as a function of experience. *Journal of Clinical Psychology*, 1957, 13, 379-382.
- LEHMANN, H. E., BAN, T. A., & DONALD M. Rating the rater. *Archives of General Psychiatry*, 1965, 13, 67-75.
- MARSDEN, G. Content-analysis studies of therapeutic interviews: 1954 to 1964. *Psychological Bulletin*, 1965, 63, 298-321.
- OSKAMP, S. The relationship of clinical experience and training methods to several criteria of clinical prediction. *Psychological Monographs*, 1968, 76(28, Whole No. 547).
- REID, D. K., & SNYDER, W. U. Experiment on "recognition of feeling" in non-directive psychotherapy. *Journal of Clinical Psychology*, 1947, 3, 128-135.
- ROGERS, C. R. (Ed.) *The therapeutic relationship and its impact*. Madison: University of Wisconsin Press, 1967.
- SHAPIRO, J. G. Relationships between expert and neophyte ratings of therapeutic conditions. *Journal of Consulting and Clinical Psychology*, 1968, 32, 87-89.
- SHOBEN, E. J., JR. Counseling. *Annual Review of Psychology*, 1956, 7, 147-172.
- TAFT, R. The ability to judge people. *Psychological Bulletin*, 1955, 52, 1-23.
- TAFT, R. Multiple methods of personality assessment. *Psychological Bulletin*, 1959, 56, 333-352.
- TAFT, R. Accuracy of empathic judgments of acquaintances and strangers. *Journal of Personality and Social Psychology*, 1966, 3, 600-604.
- TRUAX, C. B., & CARKHUFF, R. R. *Toward effective counseling and psychotherapy*. Chicago: Aldine Press, 1967.
- VARGAS, M. J. Changes in self-awareness during client-centered therapy. In C. R. Rogers & R. F. Dymond (Eds.), *Psychotherapy and personality change*. Chicago: University of Chicago Press, 1954.
- VINGOE, F. J., & ANTONOFF, S. R. Personality characteristics of good judges of others. *Journal of Counseling Psychology*, 1968, 15, 91-93.
- WATLEY, D. J. Counselor predictive skill and rated counselor effectiveness. *Personnel and Guidance Journal*, 1967, 45, 579-584.

(Received April 8, 1968)

REACTIONS OF A PSYCHOTHERAPY GROUP TO AMBIGUOUS CIRCUMSTANCES SURROUNDING THE DEATH OF A GROUP MEMBER¹

DONALD D. KIRTLEY

AND

JOSEPH M. SACKS

Fresno State College

*Veterans Administration Hospital,
Fresno, California*

This paper describes the reactions of group-therapy patients on a psychiatric ward of a general hospital to the death of a group member who died in the hospital the day before the session in question took place. The circumstances of this patient's death were ambiguous as to whether the death was due to suicide or accident. In view of the ambiguity of the situation, some members of the treatment team felt that group discussion would generate more emotional disturbance among the patients than it would reduce. However, it was concluded that discussion of death and open consideration of the feelings of guilt, anxiety, and hostility involved can be therapeutic in a group as well as in an individual setting.

From its hush-hush status as a taboo topic, particularly with people who are dying or who are closely associated with the dying, death is beginning to emerge as a respectable subject for communication. A number of studies of attitudes toward death suggest that frank and open discussion of the subject by both doctor and patient is therapeutically desirable (Feifel, 1959; Kayton & Freed, 1967; Margolis, Meyer, & Louw, 1965; Shneidman, 1963; Weisman & Hackett, 1961). The present study is concerned with the value of such discussion for a therapy group which had recently lost one of its members due to a sudden and violent death. The incident took place on the psychiatric ward of a general hospital. The body of one of the group members, a Mr. K, had been found in the basement of the hospital at the foot of a staircase from which he had apparently fallen. There was no way to determine whether the fatal fall had happened through accident or otherwise. Thus the circumstances surrounding the death were ambiguous. This study differs from those previously mentioned insofar as it deals with a therapy group rather than an individual therapeutic situation and the death discussed had occurred under ambiguous

conditions. This element of ambiguity appeared to be very important in determining the reactions now to be described.

The question of the therapeutic value in group therapy discussion of the recent death of a group member became an issue dividing the treatment team. Some staff members felt that, in view of the ambiguity of the situation, such discussion would create further emotional disturbance among the patients and that the best way to handle the matter would be simply to drop it. Others felt just as strongly that the subject should be discussed openly. This latter approach is congruent with that of Feifel (1963) who maintains that

Not only patients, but most of us feel less apart when we can share feelings and thoughts about dying and death . . . the opportunity to discuss their feelings helps subdue irrational fears and guilt sentiments associated with thoughts of death [p. 17].

While Feifel is here concerned with patients with terminal illness, we observed similar reactions in psychiatric patients following the death of a therapy-group member.

The group meeting to be discussed took place on the morning following the mishap, at which time the County Coroner's Office was still investigating the case. However, later that same day the coroner's office stated that the cause of the fall would be listed as "unknown." K had been in the hospital approximately 6 weeks and most of the patients in his group knew him personally, having

¹A brief version of this paper was presented at the Annual Meeting of the American Psychological Association in San Francisco, September 1968. Requests for reprints should be sent to Joseph M. Sacks, Veterans Administration Hospital, 2615 Clinton Avenue, Fresno, California 93703.

been on the ward with him for at least 2 weeks. On the morning following K's death the group therapy session was late in starting. Patients had to be rounded up and seemed to be dragging their feet coming in. Six patients were present, two staff psychologists (Dr. O and Dr. J), a psychologist-in-training (Dr. D), and a nursing assistant (Mr. F).

When the session began the entire room was permeated with an air of gloom and no one seemed to be willing to initiate any activity. After a few moments of silence, Dr. O commented that staff as well as patients were feeling very badly about K's death. This seemed to lift the lid of silence which had been covering the group. When the patients began talking it became apparent that most of them believed that K had committed suicide. They felt that the staff knew exactly what had happened, that there had been a suicide, that it occurred because the staff had been negligent, and that the staff was trying to cover up its mistakes by passing off the death as accidental. As one patient, a Mr. T, stated,

This whole thing stinks like shit and you guys are trying to bury it in sugar. You think that'll make you look better, but shit still stinks no matter how much sugar you dump on it, so you're not going to hide anything. You think you're protecting yourselves, but you're not. You're letting yourselves down worse than you are because everybody's going to know what happened anyhow, so lay off with all that sugar.

The group's expression of hostility toward staff was considerable, although T was far more loquacious and direct in his attacks than any other group member. Only two patients appeared to consider seriously the view that K's death had been an accident. The rest of the patients indicated an uncomfortable degree of doubt regarding staff competence and commitment. In effect they said that the staff was not genuinely interested in patient welfare for, if the staff were so interested, the tragic death of K could have been prevented. In other words, had the staff helped K more, or even if they had kept closer watch over him, he would probably still be alive. When Dr. J questioned whether one can be prevented from committing suicide when he is bound and determined to carry out the act, the patients

became highly irritated and said that the staff was merely making up excuses for itself. It should be pointed out that these patients were on an open ward. This meant that patients were free to go to their various clinic appointments, to the dining room on another floor, to occupational therapy, recreation hall, etc. They definitely wanted to keep it this way, yet they seemed to be saying that they also wanted constant surveillance and supervision by staff. The dilemma inherent in protection of the suicidal patient in this type of open therapeutic setting has been discussed by Margolis et al. (1965).

The belief that K had committed suicide was not unfounded as he had recently made an abortive attempt to shoot himself. K's reactions generally were such as to predispose him to this type of resolution of his situation. K was 59 years of age, although he looked considerably older. He was becoming increasingly worried about his health and financial problems. He had been suffering recurrent episodes of depression with anxiety for at least 5 years.

After a period of apparent improvement, K was scheduled to meet with the staff to discuss a possible discharge. K was seemingly in great conflict over whether he should or should not be discharged. He had confided to other patients that he was very anxious about leaving the hospital because he did not feel that he was as yet up to managing the frustrations his job would be imposing on him, and when he asked Dr. O for his opinion in the matter, Dr. O told him that he felt K was not yet ready to leave. However, there were numerous times when K verbalized a desire to leave anyway. On the morning of the weekly staff meeting he had apparently decided he wanted to be discharged, but this decision had yet to be discussed with the staff in conference. When K's name was called to enter the conference room, he could not be found. He was nowhere on the ward. Later he was found dead in a pool of blood lying at the foot of the stairs leading to the basement. His skull was fractured in several places and he had apparently died instantly from the injuries received in his fall.

The patients in the group, especially T, angrily continued to support the view that K

had committed suicide and seemed to resent any other interpretation, such as that the death had been accidental or simply that the circumstances leading to death were unknown, the latter actually being the most conservative and reasonable judgment in light of the facts then available. Then T challenged those staff members present by asking them whether they felt they had done all they could have done for K and the staff members replied that they had been searching their consciences and asking themselves if there was anything they had done or failed to do which might have prevented the incident. They recalled occasions on which K approached them when they had appointments with other patients. He could have felt rejected when they exchanged only brief comments with him or told him they would see him later.

Then rather suddenly group activity turned in a different direction when Dr. O asked the patients whether they felt that they had done all they might have in order to help K. Had they themselves any guilt regarding their own treatment of K? When the patients were thus confronted, their hostility began to dissipate and a considerable amount of guilt was admitted to, especially by Mr. T and Mr. S. The patients generally stated that they felt rather badly because they did know they might have been a little more helpful to K than they actually were, but S seemed to feel worse than the others, excepting T. S said that on the very morning of the mishap K had approached him in a friendly manner, apparently because he felt in need of a little companionship and conversation, but at the time S did not feel like talking; he was preoccupied with his own problems and simply wished to be left alone. Consequently, he turned his back on K and walked away from him without saying one word in answer to K's friendly advance. Furthermore, T, who up to this time had been the most vociferous in attacking the staff, stated that he had been seriously troubled about something he said to K during an earlier group session, the session during which K described his attempted suicide. On that particular occasion T expressed himself in a pessimistic and defeatist manner while K was in the midst of relating his feelings during and after the

attempt. T said, "Well, if that's how a guy feels, he ought to go ahead and do it. I can't blame you. If you feel that way about life, there's nothing to do but blow your brains out. Hell, you can't win 'em all." On the day of K's death, a short time after the death was announced, T recalled this bit of misdirected advice and became agitated and preoccupied with the thought that maybe these very words had some part to play in helping K come to the point where he could make the fatal decision. He went to see Dr. O to talk over his feelings in the matter. In the group meeting the following day T stated that he still felt a high degree of guilt on account of what he had said to K, though at this time he said that he did not think it had any real influence on K's behavior. Thus, it was possible that T's strong hostility stemmed from an acute and intense guilt reaction projected upon the staff. Such projection was characteristic of T's general behavior.

Aside from realistic concern and guilt, a great deal of anxiety seemed to underlie the hostility which had been expressed in the group. A Mr. J, who appeared to be the most openly anxious of the group members, began toward the end of the session to talk about how his mother might react to K's death if she were to learn of it. J was afraid that if he told his mother of the incident she would become highly upset, as K might have been a member of the family and she was always disturbed by the death of a family member. However, J's concern in this matter seemed to be a curious concern since, if K was related to him, it was only in a distant way and, furthermore, J's mother did not know K at all nor had she ever seen him. Nonetheless, J did appear to be disturbed and it seemed that the talk about his mother was merely a groping effort on his part to find some sort of rationalization to help bind his anxiety. Anxiety was suggested by patients' verbalizations generally. It was as though they were continually saying, "If something horrible like that could happen to K, then what might happen to me in this place? What kind of treatment am I getting anyway? Whatever K got sure didn't do him any good and maybe it even hurt him." In other words, K's mishap had shaken the pa-

tients' credibility in the efficacy of the ward's treatment program, of which previously they had probably held a higher opinion. Their new and unexpected perceptions of their situation may have produced conflict and anxiety regarding possible modes of reaction to the new experience.

By the end of the session the group still did not appear to have reached adequate closure with regard to the events leading up to K's death. Hardly any of the patients were satisfied with the "unknown" theory which was essentially the position put forth by the staff. At this point, Dr. D presented the hypothesis that K's death may have been an unconsciously motivated suicide. Even if K had died accidentally, there still may have been a "self-destructive" element at work. Everyone knew that K was still depressed and had previously thought of suicide. Even though K had recently been trying to overcome his depression, this effort may have been taking place at the conscious level only, while unconsciously self-destructive impulses may have been actively at work. K's excessive apologizing about his attempted suicide suggested a reaction formation to the original consciously felt death wish. Assuming that K did have an accident, factors such as the above could have led him to be more accident-prone than he otherwise, under more favorable conditions, would have been. Thus, the interpretation, somewhat paradoxical at first glance, but perhaps less so after further inspection, was that the mishap had been an "accidental suicide." Patients generally, as well as staff present, seemed to find this view plausible and a fair degree of closure on the matter was thus obtained. This interpretation is consonant with Shneidman's concept of subintention. In seeking to establish a more scientific classification of the modes of death than the currently prevalent one of "natural, accidental, suicidal, or homicidal," Shneidman (1964) speaks of basic attitudes toward cessation of life as "intentioned, subintentioned, and unintentioned." He relates the subintentioned orientation to instances in which the individual plays an indirect, covert, partial, or unconscious role in his own demise. Such was the thesis presented to the group by Dr. D

when toward the end of the session the group still did not appear to reach adequate closure with regard to the events leading to K's fatal fall.

On the next meeting of this group 2 days later, patients made no mention of K. K's name was not brought up until the end of the session and it was Dr. O who mentioned it. He asked the patients whether they still had feelings about K's tragedy and expressed a degree of surprise that no one had anything to say about it. Some of the patients said they had been trying not to think of the incident and were still somewhat bothered by it, but generally the group did not appear to be willing to launch into another discussion on K. Either they had sufficiently ventilated their feelings during the previous session and no longer had any need to talk them out or they were controlling their anxiety through attempting to repress the whole incident or the newspaper report concerning the coroner's judgment that the mode of the death was accidental and circumstances leading to death were simply unknown (a judgment which prior to the coroner's report they would not accept from the staff) had given them final closure on the subject. To some extent all of these factors were probably present.

The study of Kayton and Freed (1967) of the effects of a suicide in a psychiatric hospital showed strikingly similar reactions to those of our patients. In both a closed ward of disturbed patients and in an open, milieu-oriented ward, patients were angry with staff for invoking the pain of insecurity by not preventing the suicide. This reaction appeared to be aggravated on the closed ward when the usual staff-patient meeting was delayed several days. On the ward on which the suicide actually occurred, the psychiatric resident conducted a meeting with the patients that same evening. Interestingly, in this situation the physician's manifest anxiety evoked protective reactions by the patient group who reassured him that nothing could have been done to prevent the suicide. Unlike our patients, and those of the other two wards described, these patients focused on the deceased as the cause of their insecurity and utilized mechanisms of denial in resisting any further discussion about

the incident. In their therapeutic recommendations, Kayton and Freed suggest that a patient-staff meeting should be convened as rapidly as possible following such an incident, and that the group leader should supply the specific facts of the incident to the group. They conclude that "as it becomes apparent that the stress is dissipating, no further patient group ventilation is indicated unless it is expressly desired by the patients [p. 194]."

Contrary to the opinions of some staff members, open group discussion of the death did not generate any increase in emotional disturbance. If anything, the opportunity to express feelings and to approach a consensus about the cause of death, seemed to have prevented a build-up of tension in this group.

REFERENCES

- FEIFEL, H. Attitudes toward death in some normal and mentally ill populations. In, *The meaning of death*. New York: McGraw-Hill, 1959.
- FEIFEL, H. Death. In N. L. Farberow (Ed.), *Taboo topics*. New York: Atherton Press, 1963.
- KAYTON, L., & FREED, F. Effects of a suicide in a psychiatric hospital. *Archives of General Psychiatry*, 1967, 17, 187-194.
- MARGOLIS, P. M., MEYER, G. G., & LOUW, J. C. Suicidal precautions: A dilemma in the therapeutic community. *Archives of General Psychiatry*, 1965, 13, 224-231.
- SHNEIDMAN, E. S. Orientations toward death. In R. W. White (Ed.), *The study of lives*. New York: Atherton Press, 1963. (Reprinted in *International Journal of Psychiatry*, 1966, 2, 167-200.)
- WEISMAN, A. D., & HACKETT, T. P. Predilection to death. *Psychosomatic Medicine*, 1961, 23, 232-256.

(Received April 8, 1968)

PERCEPTUAL RECOGNITION IN CHRONIC AND ACUTE SCHIZOPHRENICS¹

EUGENE EBNER AND BARRY RITZLER²

Lafayette Clinic, Detroit, Michigan

Schizophrenic and normal groups were compared on their speed in recognizing the identity of pictorial stimuli which gradually came into sharper focus from an initially blurred appearance. The chronic schizophrenic group showed slower recognition than the acute schizophrenic or normal groups, but the groups did not differ in amount of prerecognition hypotheses or time of first response. The differences in perceptual recognition speed seem related to certain clinical differences between the acute and chronic schizophrenic, particularly with regard to the impact of external perceptual stimuli vs. internally generated ideation.

The nature of the cognitive disorder in schizophrenia has been extensively described and studied from many vantage points. However, the specific mechanisms and mediational processes contributing to this dysfunction are still poorly understood. Several recent conceptualizations of schizophrenia (McReynolds, 1960; Silverman, 1964) have pointed to the possibility that basic systems, or parameters of information processing, may be altered in schizophrenia in ways that predispose the schizophrenic toward his distorted and bizarre interpretations of external and internal stimuli. McReynolds (1960), for example, proposes that the schizophrenic's cognitive anomalies may result from inability to assimilate incoming perceptual input rapidly enough, with consequent disruptive anxiety and breakdown of reality-based perception and cognition. Silverman (1964) reviewed studies of various "cognitive control" dimensions and proposed a schizophrenic deficit in a "sensory input processing-ideational gating" mechanism which involves alterations in modes of information reception and evaluation. Although the specific conceptualizations offered by these writers still remain to be empirically

evaluated, they are of importance in focusing experimental interest on the operating characteristics of central perceptual-cognitive processes, *qua* process, in contrast to the more traditional attempts to identify mainly peripheral performance differences between schizophrenic and nonschizophrenic groups.

Clinical observations have provided vivid examples of the extent to which the schizophrenics' anxieties and fears of being overwhelmed are intensified and perhaps activated by the distortions of meaning which are assigned to various environmental occurrences—or, in other terms, to aberrant perceptual-cognitive information processing and categorization. Thus, it would appear that a crucial aspect of these processing systems is the stage in which meaning, or object-event identity, is assigned to perceptual input; that is, when a perceptual array is "recognized" in terms of the individual's categories of reference. Furthermore, since perceptual activity can almost always be thought of as a temporally extended sequence of events (Smith, 1958), even when the temporal interval is no longer than in tachistoscopic exposures, it seems important to study the pattern or sequence through which various final perceptual responses pass (see Flavell & Draguns, 1957, and Werner, 1956, on the "microgenetic" aspects of perception). This paper describes a study in which the perceptual recognition process in schizophrenics was investigated with the aim of determining characteristic features in the temporal pattern of processing perceptual input which might further

¹ This investigation was supported in part by United States Public Health Service Grant MH10982-01 from the National Institute of Mental Health to the senior author. Portions of this paper were presented at the American Psychological Association meeting, Washington, D. C., September 1967.

² Both authors are also at Wayne State University. Requests for reprints should be sent to Eugene Ebner, Lafayette Clinic, 951 E. Lafayette, Detroit Michigan 48207.

elucidate modes of cognitive dysfunction in schizophrenia.

Under normal conditions of perceiving, however, the steps in the temporal development of a perceptual response are difficult to isolate, partly because of the rapidity of the process and the presence of highly redundant cues. In tachistoscopic threshold procedures, it is possible, of course, to slow down and control the recognition process but only by transforming the naturally occurring continuous perceptual influx into a series of discontinuous, discrete stimulus presentations. In this study the aim was to "slow down" the perceptual recognition process by controlling the presentation of input, but, at the same time, to maintain continuous input during the experimental exposures. For these reasons a perceptual recognition technique developed by Bruner and Potter (1962, 1964) which allows the presentation of increasing amounts of stimulus information in a continuous manner was employed. In this procedure, *S* views photographic slides which are initially presented extremely out of focus, and which are continuously brought into sharper focus over a given time interval during which *S* attempts to identify the content of the slide. The variables measured were the point of correct recognition of the picture, the number of responses given prior to correct recognition, and the point of first response. These variables may be regarded as aspects of the mediating steps and transformations of data between perceptual input and response, and thus are pertinent to the activity of perceptual organizing processes. Previously, Davis and Cullen (1958) studied perceptual recognition in schizophrenic, neurotic, and normal groups using a somewhat similar procedure in which photographs were presented in increasing clarity by manipulating illumination level. They found significantly elevated recognition thresholds for schizophrenics versus neurotic and normal groups. These data, however, are based upon responses to only two photographs with unsystematically varied rate of illumination change, so that the stimulus presentations are not strictly comparable.

More recently, Draguns (1963) studied the responses of schizophrenics to perceptually

ambiguous material by requiring recognition of drawings which were blurred in varying amounts and then presented in a series from very blurred to sharp. He found that chronic schizophrenics gave earlier recognition responses and more perceptual recognition errors than acute schizophrenics or normals, and concluded that chronic schizophrenics tend toward more premature and idiosyncratic structuring of incomplete perceptual data. It should be noted, in line with the above comments regarding the discrete nature of tachistoscopic procedure, that Draguns' task also involved discrete stimulus presentations.

While the present study is addressed to similar questions as the Draguns (1963) and Davis and Cullen (1958) studies, the methodology employed here allows for more control over the perceptual information presented the *S*, while retaining the feature of continuous perceptual stimulation during each trial. Both chronic and acute schizophrenic groups were studied.

METHOD

Subjects

Chronic schizophrenic and acute schizophrenic groups were chosen from patients in a state mental hospital,^a and a training and research hospital. Patients were included only if they had a staff diagnosis of schizophrenic reaction, at least eighth-grade education, and revealed no evidence of organic factors in their clinical condition. Almost all the patients were receiving psychoactive medication which it was not possible to discontinue for purposes of the study.

Two control groups were formed, one consisting of state hospital nursing assistants and maintenance personnel. This group, however, did not seem a fully appropriate control group for the acute schizophrenic group which tended to be younger and more highly educated. Accordingly, a second control group of university students was formed. Thus, the four groups which were finally seen were as follows:

Group 1. Chronic schizophrenic: 20 male patients between 21 and 40 years of age (mean age, 33.2 years), mean educational level of 10.8 years, and mean length of hospitalization of 7.8 years.

Group 2. Acute schizophrenic: 20 male patients between 18 and 42 years (mean age, 27.4 years), mean education of 13.1 years, and mean length of hospitalization of 1.9 months.

^aAppreciation is extended to Calvin Chen, Director of Research, Northville State Hospital, Northville, Michigan, and his staff for their cooperation in providing facilities and *Ss* for this study.

Group 3. Hospital control: 20 male *Ss* between 18 and 40 years (mean age, 28.7 years), and mean education of 12.7 years.

Group 4. College control: 20 male Wayne State University students between 18 and 44 years (mean age, 21.8 years) and mean education of 13.1 years.

Apparatus

The stimuli were presented from a Kodak Carousel slide projector which was modified by attaching a gear system and motor to the shaft of the focusing mechanism on the projector. This system made it possible to control the speed of the forward or backward movement of the lens to focus the picture. A rheostat attached to the motor allowed fine adjustment of the lens movement. In this experiment, the lens moved at a constant speed from grossly out of focus to the point of sharpest focus in 120 seconds.

The stimuli consisted of ten 35-mm. color slides of the following objects: fire hydrant, female fashion model, telephone, highway, books, statue (Rodin's *Thinker*), pieces of silverware, baseball player, man and woman in a dancing pose, several carpentry tools. For each *S* the slides were presented in the above order, and were arranged in the slide tray with a cardboard blank between each slide.

Procedure

The *S* was seated at one end of a long table facing a rear projection screen immediately in front of him. He was given the following instructions:

We are trying to find out more about the way people see things and learn to recognize what they see. In order to do this, we need to have the help of many different people. That's why we are asking people in this hospital to help us with this experiment. The results of this experiment will not go on your record and will not in any way influence your stay in the hospital.

This is a screen and behind it are a slide projector and some equipment to operate it. I'm going to show some pictures on this screen. At first they will be very blurred and out of focus and you probably won't be able to tell what they are, but they will gradually get clearer and clearer until you can definitely tell what they are. What I would like you to do is to look at the picture and try to identify it as it is coming into focus. You may try to guess what it is at any time, even if you are not sure. Also, while the picture is coming into focus I would like you to describe what you see even though you can't tell what it is. You may talk about it in any way and remember that you may guess what it is at any time. Do you have any questions? I'm going to give you a practice picture first. Be sure to say all your ideas about the picture while it is coming into focus. When I say "ready," start looking at the screen and I'll put on the first picture.

As soon as each slide appeared on the screen, the projector lens began moving toward sharper focus.

The *S's* responses were tape-recorded and measured in terms of seconds elapsed from stimulus onset until the response was spoken. It was also possible to note the amount of lens movement to the time of the response. Each slide was viewed for the full 2-minute exposure time even when it had been recognized at some intermediate point in the exposure interval.

RESULTS

The data were analyzed with regard to the variables of time of correct recognition, time of first response, and number of responses given prior to the correct recognition response. In computing the statistical distributions, it became apparent that there was marked heterogeneity of variance among the groups on the various pictures, which precluded an effective factorial ANOVA across all groups and pictures simultaneously. Since variance transformations did not sufficiently reduce the heterogeneity, it was decided to separate those pictures showing the smallest variance among the groups from the remaining pictures, and to perform separate variance analyses for these two groups of pictures.⁴ The justification for this procedure is that because the pictures differed markedly in their perceptual properties and complexity as they were coming into focus, one can consider them as comprising subsets of differing perceptual difficulty. Grouping the pictures according to the amount of variation in performance which they produce would appear to be a meaningful way of recognizing these different perceptual characteristics. In this way, separate analyses were done on Pictures 1, 3, 4, 6, and Pictures 2, 5, 7, 8, 9, 10 for each dependent variable. Table 1 shows the mean performance scores of each *S* group, and Table 2 contains the results of the analysis of variance for each group.

Recognition response. The results of the analyses for both picture groups (A and B) were consistent in revealing significant ($p < .01$) differences among the four *S* groups in mean recognition time, as well as marked differences between pictures ($p < .01$) in both

⁴ The picture content of these two groups is as follows: Group A—Picture 1: fire hydrant, Picture 3: telephone, Picture 4: highway, Picture 6: statue. Group B—Picture 2: female fashion model, Picture 5: books, Picture 7: pieces of silverware, Picture 8: baseball player, Picture 9: couple dancing, Picture 10: carpentry tools.

Picture Groups A and B. This large variation among the pictures is mainly attributable to the quite varied perceptual qualities of the separate pictures. The interaction of Pictures \times S Groups reaches significance ($p < .05$) only with Pictures 1, 3, 4, 6. Individual comparisons showed that the chronic schizophrenic group differed significantly from each of the other three S groups which in turn were not appreciably different among themselves. Although the significant interaction term in Picture Group A suggests that the S groups' recognition performance should be considered in relation to specific pictures, the magnitude of this interaction effect was clearly overshadowed by the variation attributable to the diagnostic classifications and the general between-picture differences. The most noteworthy aspect of this analysis would appear to be the wholly consistent pattern that on each of the 10 pictures the chronic schizophrenic group had the longest mean recognition time.

First response time. Analysis of this variable disclosed a pattern similar to the recognition response data. The S groups differed significantly in time of first response in both picture groups ($p < .05$ for Picture Group A, $p < .01$ for Picture Group B) with individual

TABLE 1
MEAN SCORES ON RECOGNITION RESPONSE TIME,
FIRST RESPONSE TIME, AND NUMBER OF
PRERECOGNITION RESPONSES

Group	Recognition response (in seconds)		First response (in seconds)		Pre- recognition responses (number)	
	A	B	A	B	A	B
Chronic schizophrenic	94.62	95.84	30.43	36.18	3.21	2.82
Acute schizophrenic	56.70	77.92	18.59	24.80	2.82	3.73
Hospital control	64.97	83.13	18.96	20.03	3.19	4.67
College control	54.22	81.43	9.41	15.03	3.44	4.60

Note.—Entries in A columns refer to mean scores over Pictures 1, 3, 4, 6. Entries in B columns refer to mean scores over Pictures 2, 5, 7, 8, 9, 10.

comparisons indicating that the mean first response time of the chronic schizophrenic group was significantly longer than in each of the other three groups.

Prerecognition responses. There were no significant differences among the S groups in mean number of responses given prior to correct picture recognition, although the number of prior responses given to specific pictures did show significant variation ($p < .05$ for Picture Group A and $p < .01$ for Picture Group B).

TABLE 2
ANALYSIS OF VARIANCE OF RESPONSE VARIABLES

Source	df	Recognition response		First response time		Prerecognition responses	
		MS	F	MS	F	MS	F
Picture Group A							
Between Ss	79						
Diagnosis	3	27597.27	32.62**	5930.59	3.99*	5.14	.21
Error between	76	846.27		1485.53		24.75	
Within Ss	240						
Pictures	3	23442.57	37.82**	4097.53	10.75**	11.70	3.48*
Diagnosis \times Pictures	9	1309.52	2.11*	264.79	.69	2.73	.81
Error within	228	619.85					
Picture Group B							
Between Ss	79						
Diagnosis	3	7324.83	21.85**	9810.88	3.04*	90.70	2.32
Error between	76	335.12		3223.57		39.07	
Within Ss	400						
Pictures	5	6846.92	42.58**	532.06	1.50	15.39	5.59**
Diagnosis \times Pictures	15	249.21	1.55	670.30	1.89	4.74	1.72
Error within	380	160.82		355.42			

* $p < .05$.

** $p < .01$.

DISCUSSION

The attainment of correct recognition in this experimental procedure is dependent upon the continuous reception and matching of incoming perceptual data, or cues, against the internal organizations and categories in terms of which past experience is represented and stored. This cue-category matching process (Bruner, 1957) presumably eventuates in the formation of hypotheses which express the *S*'s momentary estimate of what the perceptual object is. The hypotheses are based upon some weighted combination of immediately impinging stimulus cues and attributes drawn from memory sources. It is plausible to consider that the longer the stimulus must be viewed before recognition occurs, the influence of the immediate stimulus factors in the hypotheses chosen becomes greater and more compelling as the features of the picture become sharper. With this in mind, the finding in the chronic schizophrenic group of significantly longer recognition times—which can be taken as an index of the amount of stimulus information required for accurate perceptual categorization—suggests that the chronic phase of the schizophrenic disorder may be characterized by a relatively higher threshold for assimilation of cognitively meaningful material, or in other terms, a diminished capacity for utilizing incomplete perceptual cues. The finding of significantly longer first response times in the chronic schizophrenic group would lead to a similar interpretation. Moreover, since the chronic schizophrenics did not give fewer pre-recognition responses than the other groups, it would seem difficult to account for their delayed recognition times on the basis of such factors as general response inhibition, apathy, or inability to attend to the stimuli. While it is conceivable that the slower recognition time of the chronic schizophrenic group might be partially attributed to lessened contact with the objects in the slides due to their institutionalization, there is little ground for considering this to be a crucial variable. In fact, one might consider that although there may be less actual contact with these materials in the hospital environment, the great amount of time that these patients spend

watching television provides them with a variety of object stimulation. It is also important to note that observation of the chronic schizophrenic *Ss* while doing the task seemed to indicate a high degree of interest, involvement, and sustained attention on their part. Consequently, it appears consistent and valid to interpret the differences in recognition response time, first response time, and number of prerecognition responses as arising from variations in the manner in which ambiguous perceptual information is treated by the *S* groups.

The significant main effect for pictures in both Picture Groups A and B indicates that perceptual recognition time varied significantly among the pictures in each group, among all *S* groups. This variation is to be expected in view of the marked differences in perceptual properties of the 10 pictures as they are coming into focus, in such factors as colors, contours, and spatial distribution of the elements of the picture. However, the various picture differences did not seem to exert any strongly selective effect upon the *S* groups, as seen by the generally nonsignificant interaction terms.

It is of interest that the acute and chronic schizophrenic groups fell in marked contrast to each other on recognition time, with the acutes showing the most rapid recognition of all groups on six of the 10 stimulus pictures. This pattern, although not entirely consistent across all stimuli, warrants the suggestion that the acute schizophrenic can more readily assimilate and match incoming perceptual cues to reference categories, while the chronic schizophrenic cannot as efficiently transform incomplete perceptual data into accurate object representations.

Consideration of certain clinical differences between acute and chronic schizophrenics may make these differences in perceptual performance more understandable. Specifically, in the acute stage the schizophrenic is typically in a state of intense anxiety as he experiences the new and alarming phenomena of the psychotic disorganization, with greater fluidity of perception and loss of stability of the previously familiar and reliable object world. In this profoundly frightening condition, the

individual is still trying to regain control of the object world and is sensitively attuned and responsive to perceptual input, which must be immediately evaluated in terms of its aversive or nonthreatening qualities. In contrast, the chronic schizophrenic has typically reached a somewhat stable resolution in which anxiety is decreased, bizarre perceptual and ideational material is familiar, well established, and no longer terrifying, and the "patient seems to have accepted his illness [Arieti, 1955, p. 380]" and turned away from external reality into fantasy preoccupation. From these clinical considerations, it would be expected that the acute schizophrenic—still trying to master external reality, and therefore still more attuned to external input, would show shorter perceptual recognition times, and that the chronic schizophrenic, because of the preoccupation with internally generated ideation would require more intense or compelling external stimulation to bring about perceptual recognition. This formulation does not necessarily imply that the chronic schizophrenic would show general deficits in thresholds for awareness of less cognitive sensory input (e.g., light, sound, touch) but it does imply that the *processing* of cognitive input—the evaluation and matching of cues with cognitive categories—seems differentially affected in the two phases of schizophrenia. This formulation of the perceptual recognition deficit in the chronic schizophrenic bears some similarity to the "schizophrenic defense mechanism" of "sensory input processing—ideational gating" proposed by Silverman (1964) in which the schizophrenic tends to "gate out" or minimally respond to disturbing ideational input while being able to deal more adequately with sensory input.

While it would seem desirable to compare the findings of this study with Draguns' (1964) study of perceptual recognition in schizophrenics, particularly since he reports earlier recognition judgments by chronic schizophrenics than acute schizophrenics or normals, several methodological differences render comparison difficult. For one thing, Draguns' "recognition response" includes both correct and incorrect recognition statements

so that his measure is essentially one of initial responsivity regardless of perceptual accuracy. Further, his conceptualization principally concerns the schizophrenics' "erroneous idiosyncratic structuring of ambiguous materials." His findings were that "chronic schizophrenics made earlier perceptual judgments as well as more perceptual recognition errors," which suggested "a preponderance of endogenous factors in such judgments over exogenous object determined ones [p. 26]." However, with regard to the point in the series at which veridical object recognition would have been attained, his chronic schizophrenic group may have required more data than the other groups. It is not possible to directly determine this from his data since he allowed only one recognition response per picture series. In the present study, the perceptual recognition response refers only to veridical recognition of the object, in which the "exogenous object determined factors" take precedence over idiosyncratic endogenous ones. Further, the differing modes of stimulus presentation—in Draguns' study, discrete trials, and in this study, a continuously viewed and changing perceptual array—present somewhat different information-processing situations. It would be of interest to employ Draguns' stimulus materials in the continuous presentation procedure to assess the importance of this difference.

The results of this experiment suggest that the particular methodology employed is well suited to the study of the temporal processing of perceptual information. However, further refinements are necessary to more clearly identify parametric aspects of the perceptual recognition process. For example, it would be highly useful to have some quantitative measure of the perceptual information available at various stages of focus, perhaps with certain standard pictures. Also, it would be of great interest to know the manner in which the stimulus display is scanned prior to correct recognition by studying eye-movement patterns. Studies such as these would afford a more detailed examination of the component processes in perceptual recognition and of the mechanisms through which the schizophrenics' anomalous cognitions may operate.

REFERENCES

- BRUNER, J. S. On perceptual readiness. *Psychological Review*, 1957, 64, 123-152.
- BRUNER, J. S., & POTTER, M. C. An exploration of visual recognition. Paper presented at Yale University, May 1962.
- BRUNER, J. S., & POTTER, M. C. Interference in visual recognition. *Science*, 1964, 114, 424-425.
- DAVIS, D. R., & CULLEN, J. H. Disorganization of perception in neurosis and psychosis. *American Journal of Psychology*, 1958, 71, 229-237.
- DRAGUNS, J. G. Responses to cognitive and perceptual ambiguity in chronic and acute schizophrenics. *Journal of Abnormal and Social Psychology*, 1963, 66, 24-30.
- FLAVELL, J. H., & DRAGUNS, J. G. A microgenetic approach to perception and thought. *Psychological Bulletin*, 1957, 54, 197-217.
- MCREYNOLDS, P. Anxiety, perception, and schizophrenia. In D. D. Jackson (Ed.), *The etiology of schizophrenia*. New York: Basic Books, 1960.
- SILVERMAN, J. The problem of attention in research and theory in schizophrenia. *Psychological Review*, 1964, 71, 352-379.
- SMITH, G. Visual perception: An event over time. *Psychological Review*, 1957, 64, 306-313.
- WERNER, H. Microgenesis and aphasia. *Journal of Abnormal and Social Psychology*, 1956, 52, 347-353.

(Received April 12, 1968)

PSYCHOTHERAPY PROCESS AS A FUNCTION OF THE TIME SEGMENT SAMPLED

NORMAN J. KARL¹ AND NORMAN ABELES

Michigan State University

A content analysis of individual psychotherapy tapes of 25 female college students was employed to determine the relative frequencies of certain response measures during the first, second, third, fourth, and fifth 10-minute segments of the therapeutic hour. An attempt was made to test the assumption made by many investigators that the occurrence of a given phenomenon is evenly distributed over the course of therapy. The results indicate that such an assumption is grossly inaccurate—given variables were found to appear consistently more frequently in certain segments of the hour than they do in others. It is suggested that examining the variable where it occurs most often will increase the power of future investigations.

Several authors (cf. Bordin, Cutler, Dittmann, Harway, Raush, & Rigler, 1954; Kiesler, 1966) have pointed to the methodological factors which need to be considered in psychotherapy process research. Among others, sampling decisions ought to be tied to the theoretical framework to which one's measures are related, the size of the segment to be sampled must be determined, the location of the sample taken from the therapy hour has to be chosen, and the rating systems used should ideally represent unidimensional scales. Most of these questions need to be put to empirical test, and decisions regarding the validity of the assumptions underlying much psychotherapy research must be based on data rather than left at the level of speculation.

A number of studies of psychotherapy process have utilized an analysis of tape-recorded samples of therapeutic sessions (Kiesler, Mathieu, & Klein, 1964; Tomlinson & Hart, 1962; Truax & Carkhuff, 1967; Van der Veen, 1967). These investigators have employed the selection of random time samples from a number of therapy sessions. This method tends to assume that the form of a given phenomenon is essentially the same during the first third of the hour (or the first 4 minutes) as it is during the last third of the hour (or the last 4 minutes). While it is often assumed that random sampling provides the

least biased data, one can legitimately raise the question whether these techniques actually "wash out" true differences. Van der Veen (1967), in fact, suggests the possibility that segment selection may have been in part responsible for his inability to replicate results by previous investigators.

Kiesler, Klein, and Mathieu (1965) provide strong evidence to support the notion that the time segment sampled represents a critical variable for the content obtained. Further, such variation seemed to be influenced, in their investigation, by the particular diagnostic group to which an S belonged. Some interesting issues are raised as a result of that investigation:

1. Are assumptions regarding variance as a function of segment sampling applicable to a variety of process variables?

2. Can such variance be demonstrated over the course of the entire therapy process?

The essential questions under consideration are whether or not changes occur in the form of the therapeutic process, when client and therapist populations and process variables differ from those employed in other investigations. Are certain response categories more typically employed (across all clients and over the course of therapy) during various phases of the therapeutic hour? At issue is the tenability of the assumption concerning the invariance of the process for a given response characteristic during or across the therapeutic interview.

¹Now at Counseling Center, University of Maryland. Requests for reprints should be sent to Norman Karl, Counseling Center, University of Maryland, College Park, Maryland.

METHOD

Subjects

The clients selected for inclusion in the present study were all female undergraduate students who completed a minimum of six therapy interviews at the Michigan State University Counseling Center. They were self-referred clients seeking help with personal-social problems. An intake interview determined that a program of therapy was warranted. A total of 25 Ss comprised the present sample.

The Ss included in the study represent 25 of the total of 54 randomly selected clients who were asked to participate in the research program being conducted at the Counseling Center. A sound-recording tape library was developed on the psychotherapies of these 54 clients, and it is from this library that the current sample was drawn.

Therapist

Clients were assigned to therapists on the basis of matching availability of hours. Meetings were generally conducted weekly, except in the case of practicum therapists, who met twice a week with their clients. Final assignment of a case depended upon the therapist's willingness to work with the client. Such judgments were made on the basis of intake interview data and the personal impressions of the intake interviewer.

The psychotherapists in the sample represented several experience levels: six staff counselors were PhD counseling or clinical psychologists with 3 to 20 years of psychotherapy experience; 14 interns were advanced counseling or clinical doctoral candidates with a minimum of 1 year of intensive psychotherapy supervision; five practicum students represented a group of less advanced graduate students who were involved with their first cases as part of an introductory practicum course.

Sampling Procedures

All therapy tapes for each of the 25 Ss were included in the present investigation. The sample represented a total of 330 tapes. Each tape was randomly assigned to be sampled during its first, second, third, fourth, or fifth 10-minute portion. A content analysis was performed on the 330 10-minute tape segments.

Coding Procedure

The content analysis systems employed derive, in part, from two sources: (a) the method developed by Bandura, Lipsher, and Miller (1960), and modified by Winder, Ahmad, Bandura, and Rau (1962), was applied to measure client "Dependency" and "Hostility" response categories, and therapist "Approach" and "Avoidance" behavior; and (b) Lennard and Bernstein's (1960) procedure was employed to identify client and

therapist "Primary" and "Secondary" role systems, and "Descriptive," "Evaluative," and "Prescriptive" therapist statements.

Description of Terms

Dependency statements include expressions of need to depend on others; allowing others to assume the initiative; to be told what to do; to be "cured" by an external source; description of dependent, approval seeking behavior; dependent agreement and nurturant-accepting behavior.

Hostility statements include description or expression of unfavorable, critical or sarcastic remarks; oppositional, antagonistic, and argumentative attitudes; expressions of dislike, disagreement, resentment, resistance, irritation, annoyance, and anger; expressions of aggressive domination.

Therapist Approach responses represent any verbalizations which seem designed to elicit further expression or elaboration of client verbalizations. This category includes therapist endeavors such as exploration, reflection, labeling, interpretation, support, and information giving.

Therapist Avoidance responses include any verbalization which seems designed to inhibit, discourage, or divert further expression of client statements. Subsumed under this category are therapist disapproval, topic transition, ignoring a client statement, mislabeling, and inappropriate silence.

Primary role system statements refer to verbalizations in which the client or therapist allude to their respective roles during the therapeutic process and to the purposes, goals, and accomplishments of treatment.

Secondary role system statements include verbalizations pertinent to the interpersonal interaction between the client and therapist. Transference phenomena are encountered here.

Therapist Descriptive propositions seek or offer information. They similarly offer or ask for clarification, orientation, or reiteration.

Evaluative therapist statements convey appraisal, value judgments, and interpretations. They inquire about or offer opinion, expression of affect, or analysis.

Prescriptive propositions occur where the therapist seeks or conveys directives, including giving or asking for suggestions.

The above-mentioned categories represent discrete measures. In other words, a unit either fit a particular response category or it did not; such determinations and examples are provided in more extensive scoring manuals (Bandura et al., 1960; Kopplin, 1965; Lennard & Bernstein, 1960; Winder et al., 1962).

Scoring

The basic model subsumed under both scoring procedures is that the focus is turned to the interaction sequence. Scoring is as follows: a client statement, a therapist response, and the subsequent response on the part of the client. By definition, a scoring unit is represented by the entire verbal content of a speaker with the preceding and succeeding verbalizations of the other person acting as the boundary for that unit.

The scoring sequence used may be summarized as follows: The judge listened to a client verbalization and classified it first according to the Dependency, Hostility, or Other content response category. Then the judge scored the client statements according to the Primary or Secondary role system involved, if any. Subsequently, therapist responses were scored as Approach or Avoidance. The therapist's verbalization was similarly scored for presence of Primary or Secondary role statements, and then according to the type of proposition employed (i.e., Descriptive, Evaluative, or Prescriptive).

Scoring Reliability

Two raters scored response categories—the senior author, Judge A, and a reliability scorer, Judge B, a graduate student in clinical psychology. The two judges initially scored several tapes together in order to train themselves in the use of the system. This procedure obtained until the raters achieved an .80-.85 product-moment correlation coefficient, which took about a week of training time. Training tapes were not included in the data analyzed as part of the study.

The subsequent sample on which reliability was determined included approximately one-half of the total. The first 75 and last 74 tapes were rated by both scorers in an attempt to determine whether any consistent changes in

TABLE 1
INTERJUDGE RELIABILITY COEFFICIENTS AND THE
STANDARD SCORES AND p VALUES ASSOCIATED
WITH THE DIFFERENCES BETWEEN A PAIR
OF CORRELATIONS

Measure	Correlation coefficients ^a		Z_d	p
	First 75 tape segments	Last 74 tape segments		
Dependency	.83	.80	.53	.56
Hostility	.92	.89	1.00	.31
Client Other	.98	.98	—	—
Client Primary	.91	.96	2.51	.01
Client Secondary	.84	.84	—	—
Client Other Role	.99	.98	2.09	.04
Approach	.99	.99	—	—
Avoidance	.79	.83	.64	.52
Therapist Primary	.93	.97	2.60	.01
Therapist Secondary	.92	.95	1.40	.16
Therapist Other Role	.98	.97	1.23	.22
Descriptive	.91	.93	.78	.43
Evaluative	.98	.97	1.23	.22
Prescriptive	.81	.74	1.06	.29

^a All coefficients, $p < .002$.

Rater A's scoring occurred over the course of coding. Judge A rated the remainder of the 330 10-minute segments. The necessity of establishing reliability estimates for scoring units was eliminated since Judge A indicated to Judge B where a particular client or therapist unit began by transcribing the first few words of each sequence to be scored. Judge B stayed within these confines when scoring the client or therapist response.

Interjudge coefficients and the significance of the differences between pairs of correlations are presented in Table 1. The table entries represent the correlations of the summated client-therapist-client units in a given 10-minute segment.

It is evident that the scoring categories represent highly reliable measures. In most cases where reliability was lower, this fact could be attributed to the relatively smaller frequencies observed in the particular category under consideration. Nevertheless, interjudge reliability coefficients estimated for both the first 75 and the last 74 segments sampled were all highly significant with $p < .002$ for all measures. Similarly, the Z scores obtained for

TABLE 2

TOTAL FREQUENCIES FOR MEASURES GROUPED BY TAPE SEGMENT SAMPLED

Measure	Tape segment					
	1	2	3	4	5	χ^2
Dependency	106	96	103	123	149	15.42**
Hostility	78	106	110	110	85	9.50*
Client Primary	142	89	96	58	156	59.82***
Client Secondary	16	27	22	16	29	6.59
Approach	1195	1207	1203	1197	1187	.20
Avoidance	21	9	14	19	30	12.95*
Therapist Primary	166	164	124	98	254	86.18***
Therapist Secondary	24	62	56	63	59	19.98***
Descriptive	291	168	142	159	217	74.29***
Evaluative	891	1006	1044	1016	886	23.05***
Prescriptive	35	43	30	42	114	89.70***

* $p < .05$.** $p < .01$.*** $p < .001$.

difference between two coefficients for a given measure indicate the high degree of consistency over time in the categorization of verbalizations. While some of the Z_d scores achieved statistical significance, which would suggest that a change in scoring procedures between the two judges did in fact occur over time, it can be seen from Table 1 that these differences occurred where correlations were so high, that a difference in r of .01 or .02 was critical enough to yield a significant Z_d .

Since a maximum of two judges was employed the question of possible interconfounding of the response categories becomes an important consideration. Kiesler (1966) indicates that process scales ought to tap single dimensions. He further suggests that low inter-rater reliability generally reflects the presence of several dimensions, and that attempts should be made to establish face validity as a further measure of unidimensionality.

The classes of Dependency and Hostility seem to meet some of the minimum requirements of face validity. Kopplin (1965) indicates that these two categories represent styles of interpersonal interactions which appear in common to most major theories of personality and psychotherapy. These concepts are similarly used on a common-sense basis, such that the lay person would classify as Hostility the statement, "that man was angry or resentful."

The evidence in terms of reliability reports

of other investigators seems to further establish the unidimensionality of the response categories. Winder et al. (1962), using two sets of two judges each, report the following respective reliabilities—Hostility, .89 and .83; Dependency, .91 and .92; Approach, .83 and .98; and Avoidance, .44 and .81. Alexander (1967), employing three raters, found the following coefficients—Hostility, .88; Dependency, .80; Approach, .97; Avoidance, .44; Client Primary, .87; Client Secondary, .84; Therapist Primary, .89; and Therapist Secondary, .86.

While reliability data from other studies does not exist for the Descriptive, Evaluative, and Prescriptive categories, the high correlations obtained in the present study again seem to point toward independent categories.

The frequency with which each response measure appears during any given 10-minute segment is presented in Table 2. The table entries include the sums of the response categories found in each 330 10-minute segment. Thus, Table 2 indicates the total frequencies of a given response for all clients over the entire course of their therapies.

It is evident that major differences occurred in the relative frequencies for a given response category, on the basis of the segment of the therapeutic hour under investigation. Dependency statements tended to be manifested most frequently in the last segment of the

therapy hour and least often during the middle portion. Conversely, the middle sessions appeared to have the greatest degree of Hostility statements, with the least amount appearing in the initial and final stages. Client Primary role statements occurred more frequently during the next-to-last segment. No significant differences obtained for either the Client Secondary role statements or the therapist Avoidance categories. Therapist Primary role statements did, however, show a steady decline through the fourth 10-minute segment with a sharp rise during the last 10 minutes. Therapist Secondary role communications, on the other hand, were least frequently expressed during the first 10-minute portion of the hour, and then increased and leveled off to a relatively stable incidence of occurrence for the remainder of the session. Therapist Descriptive propositions appeared with the greatest frequency during the first and last portions of the hour, with a lower stable function during the middle part. Evaluative propositions appeared in a direct inverse relationship to Descriptive propositions. Prescriptive propositions seemed to occur with equal frequency during the first four parts of the session, and increased sharply to reach their peak during the final phase.

The strength of the fluctuations during the hour was a rather marked and consistent one, a phenomenon which was maintained across several measures.

DISCUSSION

The sampling technique chosen was specifically designed to determine whether the time segment-process phenomenon reported by Kiesler et al. (1965) could be demonstrated when all therapy sessions were sampled. The former investigators sampled five consecutive 10-minute segments from the same therapy hour for each client. While they did include early and late interviews in their design, they did not study the entire therapy process. The present study permitted such analysis, and points to the pervasiveness of the function. Had the sampling included a single interview for each client, the ability to generalize to the entire therapy process, for this sample, would have been severely limited.

The results clearly reflect a marked tendency for certain segments of therapy to be characterized more by some content categories than others. Confirmation is presented for the assertion that time must be either controlled or considered when studying psychotherapy process.

The strength of the effect of time on process is impressive. While the present investigation uses a different client population (limited to females) and different measures from those which Kiesler et al. (1965) employed, essential confirmation of their results obtained.

One caution in interpreting these findings was suggested by the work of Kiesler et al. (1965), which indicated differential effects among different diagnostic groups. They reported the greatest effect of time on the process variable within the neurotic group.

The sample under investigation in the present study was compared to a normative group of Michigan State University females. The number of deviant MMPI scales was greater for Ss in psychotherapy ($\bar{X}_1 = 2.00$, $\bar{X}_2 = 1.03$, $t = 3.59$, $df = 133$, $p < .01$). No further attempt at differential diagnosis was made. Generalizations, therefore, should be limited to a group of female college undergraduates seeking psychotherapy.

Similarly, generalizations beyond the particular process variables dealt with in this study (i.e., Dependency, Hostility, etc.) should be made with considerable caution. It is important to emphasize that the current set of results are related only to the process variables under investigation. Replicability of these findings with other clients and other therapists remains to be put to the test, as does the generality of these findings with other response measures. It should be noted, however, that the present investigation demonstrates the applicability of the time-process function to variables emerging from theoretical frameworks different from those used by other investigators.

While evidence is presented that the time segment sampled affects the process, no answer is given to the question of what elements are to go into the decision of choosing which segments should be investigated. Kiesler (1966) maintains that such determinations must be

made in relation to the theoretical framework in which one is working. Where Rogerian measures, for example, are involved, random sampling may be more justified since the pervasiveness of the phenomena is emphasized. Freudian variables, on the other hand, might require sampling techniques more intentionally focused on specific problem areas.

Given that the above-mentioned theoretical implications are included in an experimental design, the present findings suggest some important considerations for further sampling techniques. If, for example, a particular investigation is primarily concerned with dependency phenomena, the characteristic mode of operation would be to sample a portion of the total tapes available for N clients. This procedure apparently results in a partial exploration, at best, since the entire process (i.e., all interviews) of therapy is not being sampled for a given client.

While random selection of time points within the hour would still be valid if the number of samples extracted was large, maximum efficiency would seem to be obtained by sampling the segment in which the variable (e.g., "Dependency") appears in its most potent form. Such a timesaving procedure would possibly permit all therapy tapes for a given client to be included in the sample, assuming that such functions are found to obtain for other investigations.

Pilot data may have to establish, for each investigation, the portion of the therapeutic hour which most dramatically characterizes the phenomenon under consideration. Once such determination has been achieved, a more thorough detailed analysis of the variable is possible through selection of the segment in which it is most powerfully exhibited.

REFERENCES

- ALEXANDER, J. F. Perspectives of psychotherapy process: Dependency, interpersonal relationships, and sex differences. Unpublished doctoral dissertation, Michigan State University, 1967.
- BANDURA, A., LIPSHER, D. H., & MILLER, P. E. Psychotherapists' approach-avoidance reaction to patients' expression of hostility. *Journal of Consulting Psychology*, 1960, 24, 1-8.
- BORDIN, E. S., CUTLER, R. L., DITTMANN, A. T., HARWAY, N. I., RAUSE, H. L., & RIGLER, D. Measurement problems in process research on psychotherapy. *Journal of Consulting Psychology*, 1954, 18, 79-82.
- KIESLER, D. J. Basic methodological issues implicit in psychotherapy process research. *American Journal of Psychotherapy*, 1966, 20, 135-155.
- KIESLER, D. J., KLEIN, M. H., & MATHIEU, P. L. Sampling from the recorded therapy interview: The problem of segment location. *Journal of Consulting Psychology*, 1965, 29, 337-344.
- KIESLER, D. J., MATHIEU, P. L., & KLEIN, M. H. Sampling from the recorded therapy interview: A comparative study of different segment lengths. *Journal of Consulting Psychology*, 1964, 28, 349-357.
- KOPFLIN, D. A. Eliciting responses in client-therapist interaction: A content analysis of initial psychotherapeutic interviews. Unpublished doctoral dissertation, Michigan State University, 1965.
- LENNARD, H. L., & BERNSTEIN, A. *The anatomy of psychotherapy*. New York: Columbia University Press, 1960.
- TOMLINSON, T. M., & HART, J. T., JR. A validation study of the process scale. *Journal of Consulting Psychology*, 1962, 26, 74-78.
- TRUAX, C. B., & CARKHUFF, R. R. *Toward effective counseling and psychotherapy: Training and practice*. Chicago: Aldine, 1967.
- VAN DER VEEN, F. Basic elements in the process of psychotherapy: A research study. *Journal of Consulting Psychology*, 1967, 31, 295-303.
- WINDER, C. L., AHMAD, F. Z., BANDURA, A., & RAU, L. C. Dependency of patients, psychotherapists' responses, and aspects of psychotherapy. *Journal of Consulting Psychology*, 1962, 26, 129-134.

(Received April 15, 1968)

SELF-DISCLOSURE AND INTERPERSONAL FUNCTIONING¹

CHARLES F. HALVERSON, JR.

AND

ROY E. SHORE

*Child Research Branch
National Institute of Mental Health*

*Center for Psychological Studies
Educational Testing Service*

Fifty-three Peace Corps trainees were administered a self-disclosure questionnaire during a pretraining assessment program. Trainees who demonstrated a readiness to confide personal information to others were found to be more well liked by the other trainees and the training staff after 6 weeks of training than trainees who were reluctant to disclose personal information. Hypotheses concerning the relations between self-disclosure, integrative complexity, and authoritarianism were supported. The accessible trainees tended to be more integratively complex and less authoritarian than the inaccessible trainees. Implications of self-disclosure for interpersonal openness and effectiveness are discussed.

Social accessibility or self-disclosure refers to the process of communicating information about oneself to other persons. The readiness to confide personal information has been shown to contribute to the development of social relationships. Jourard and Lasakow (1958), and Jourard (1959) found that the degree of liking another person is correlated with the amount of self-disclosure to that person, while Jourard and Landsman (1960) indicate that receiving self-disclosure inputs from another person leads to greater disclosure toward that person. Newcomb (1961) has also noted that the exchange of personal information leads to the establishment of friendships.

Self-disclosure is also related to behavioral measures of interpersonal competence and openness. Colson,² Frankfurt (1965), and Taylor (1965) have all found that persons reporting high self-disclosure revealed more about themselves in social interaction situations than low scorers. Jourard (1961) found that nursing students who scored high on a self-disclosure questionnaire tended to be rated a year later as high in the "ability to

establish and maintain a communicative relationship with patients" as well as showing a high degree of openness with the nursing faculty. Presumably persons who are socially open to others are seen as more interpersonally competent than individuals who are guarded and closed in their personal affairs.

Although other research has been carried out on self-disclosure, the two groups of studies above appear to be the main lines of direct construct validity evidence so far. The present study had three aims: (a) to extend the construct validity of self-disclosure further by placing it in a larger theoretical context, or "nomological network" (Cronbach & Meehl, 1955), (b) to utilize the "nomological network" to predict further aspects of interpersonal competence, and (c) to test directly whether reported self-disclosure is related to being liked by others.

Concerning the first aim, that of placing social accessibility in a larger theoretical context, the theories of authoritarianism (Adorno, Frenkel-Brunswick, Levinson, & Sanford, 1950) and conceptual systems (Harvey, Hunt, & Schroder, 1961; Schroder, Driver, & Streufert, 1967) seem particularly relevant, as does Tomkins' (1964) notion of good-bad "polarity."

The theory of authoritarianism (Adorno et al., 1950) implies that nonauthoritarians are "trusting and open" in their interpersonal relations, while highly authoritarian persons manifest distrust and suspicion of others. Since authoritarian persons are purported to view people as threatening, they should be less

¹ Based on a paper read at the Eastern Psychological Association meetings in New York, N. Y., April 1966. This research was conducted at Syracuse University and was supported by funds from the Peace Corps. The authors wish to thank David E. Hunt and Harold M. Schroder for their cooperation.

² Requests for reprints should be sent to Charles F. Halverson, Jr., National Institute of Mental Health—Building 15-K, 9000 Rockville Pike, Bethesda, Maryland 20014.

³ W. N. Colson. Self-disclosure as a function of social approval. Unpublished manuscript, 1965.

disclosing of personal information than non-authoritarians. In other words, authoritarian individuals should rate themselves as less socially accessible than nonauthoritarian individuals.

Conceptual systems theory (Harvey et al., 1961; Schroder et al., 1967) is also relevant to self-disclosure. This theory postulates that cognitive functioning varies along a dimension of information-processing ability. Interpersonal cognitive functioning at low levels of complexity is characterized by a relatively simple and inflexible view of others, while more complex functioning represents a more flexible and open interpersonal orientation. Because their simple, inflexible structures for processing interpersonal information are inadequate for dealing with many interpersonal situations, low-complexity persons presumably often react in a defensive manner. More complex persons, whose information processing is adequate to deal with most interpersonal situations, can relate in a more open, informational manner. Therefore, a positive relation between conceptual complexity and interpersonal competence and openness was expected. Relative to conceptually simple individuals, complex individuals should rate themselves as more socially accessible.

Tomkins (1964) designed a "Polarity Scale" to measure the degree to which people believe that human nature is basically good or evil. If a person views human nature, and thereby presumably fellow persons, as basically evil, he will probably behave in a more guarded and defensive manner towards others. It was predicted that Polarity (with "man is basically evil" scored high) would be negatively related to social accessibility.

The above discussion leads to the second aim of the study, that of predicting behavioral flexibility from self-disclosure. If it is assumed that conceptual complexity, authoritarianism, and polarity of beliefs play integral roles in the development of social accessibility, one would expect socially accessible persons to behave not only more openly, as Jourard (1961) found, but also more flexibly and adaptively. Thus it was predicted that self-disclosure would be positively related to behavioral ratings of "interpersonal flexibility" and "general adaptability."

As to the third aim, although research has shown that confiding personal information plays a role in the development of social relationships, there has not been a direct test of whether individual differences in self-disclosure are related to being liked by others. In addressing that question it seemed important to distinguish the formation of initial impressions from the development of stable social relationships, for the same variables may not be important in both processes. Social accessibility would seem to play an important role in the latter, but its role in impression formation seems ambiguous: Reluctance to disclose personal information may not necessarily be noticed or salient in primary impression formation. It was predicted that social accessibility would be positively related to a sociometric index taken after the group had had extensive interactions, but no prediction was made concerning the same index taken after only a brief interaction period.

The Peace Corps training program, in which this study was undertaken, was well suited for gathering the sociometric and other interpersonal behavior data necessary to test the hypotheses. The Peace Corps trainees (PCTs) arrived at the training site from all parts of the country, and, except for a few married couples, did not know each other. However, this soon changed. The PCTs spent the next 12 weeks as a very compact and closely-knit group, being together 12 to 14 hours a day, eating, attending classes, and utilizing their little free time together. This setting provided an opportunity to see who would come to be the best liked trainees in the group, and who would be judged the most interpersonally competent and flexible by the assessment and training staff.

METHOD

Subjects

The Ss were 53 PCTs being trained for teaching in East Africa and all were college graduates. Each of the PCTs participated in a 2-day pretraining assessment program in which they were administered the experimental materials.

Measures

Self-disclosure scale. The self-disclosure scale was a modified form of the one used by Jourard and Lasakow (1958). The scale consisted of 30 state-

ments of information about the self such as, "What I think and feel about religion: my personal religious views," "All present sources of income-wages, fees, allowance, etc.," and "My views on communism." The trainees were told to read each item and then indicate the extent that they had talked, or might talk about each item to three target persons: (a) a stranger, (b) an acquaintance, and (c) their best friend. Thus the PCTs made three ratings for each item. They gave a 0 when they would disclose nothing about the particular item to that target person, a 1 if they would talk in general terms, and a 2 if they would talk in full and complete detail about the item. These numerical entries for each item were summed, yielding totals which constituted the self-disclosure score.⁴

Peer nomination. After the 2 days of pretraining assessment (during which they had only minimal free time to interact), and again after 6 weeks of training, each trainee evaluated the other trainees with a sociometric inventory. Specifically, each trainee was asked to indicate the person or persons he would most like to be assigned with in his overseas station. The trainee could write down as few or as many of his peers as he wanted. This peer nomination was quite involving for the trainees because they viewed this procedure as having an actual bearing on the assignment of trainees. They each spent some time and considerable effort in making their choices. The number of positive choices received by each trainee constituted his peer nomination score on both occasions.

Assessment board ratings. The assessment and training staff met after 6 weeks of training in what was termed the midboard evaluations. Each trainee's progress in the program was assessed, and a decision made as to whether each trainee was qualified to continue training. On the basis of aptitude tests, class performance, psychiatric interviews, and training staff evaluations, the midboard assessment team rated each trainee on several 5-point scales relative to suitability for Peace Corps service. The ratings were not contaminated by the trainees' social accessibility scores, as they were not available to the raters.

Relevant to the present investigation were the mid-board ratings of "interpersonal flexibility" and "general adaptability." The raters were provided with definitions of the scales. The definition for the "interpersonal flexibility" scale stressed rating how well the trainee was able to cope with interpersonal situations, while the "general adaptability" rating was to indicate how well he could cope with problem-solving situations and stress. The assessment board arrived at a consensual rating for each trainee for each scale.

Authoritarianism. A revised 100-item version of the California F Scale (Stern, Christie, Lane, Sanford, & Webster, 1960) was employed. The revised form was

designed to tap personality rather than ideological aspects of authoritarian functioning and had several types of item reversals built in to minimize the contribution of social desirability and acquiescence response sets and to capture adequately several varieties of nonauthoritarianism.

Conceptual complexity. Conceptual complexity was measured by the Paragraph Completion Inventory (Schroder et al., 1967), a semiprojective test which requires Ss to write several sentences in reaction to and resolution of situations which imply uncertainty, structure, and conflict. In scoring the completions, the raters evaluated the number of alternative resolutions expressed or implied by the completion, as well as the complexity of the rules which S used in producing the response. The completions were scored along a 7-point scale representing a continuum from low to high complexity (see Schroder et al., 1967, Appendix 2, for scoring manual) by two experienced scorers, who had in the past consistently shown interjudge reliabilities in the .90s. The score for each trainee consisted of the average of the scores assigned to the six responses by the two judges.

Polarity Scale. The Polarity Scale (Tomkins, 1964) consisted of 56 forced-choice format items administered with the instructions used by Tomkins that S could check either, both, or neither of the two alternatives for each item. The forced-choice format minimized the possibility of responding to the alternatives on the basis of greater social desirability of one of the items. A trainee's score was determined by dividing the number of "human nature is basically evil" choices checked by the total number of choices checked.

Aptitude tests. The School and College Aptitude Test (SCAT) verbal scores and General Aptitude Test scores were available on all Ss, and were used in this study to check on whether the finding of hypothesized correlations among other variables could be attributed to a joint relationship to intelligence. This check can also serve as one indication of the discriminant validity of the social accessibility measure.

RESULTS AND DISCUSSION

The product-moment correlations between self-disclosure and the other measures are shown in Table 1. As can be seen in Table 1 nearly all the predicted correlations with self-disclosure were significantly different from zero in the expected directions.

These data offer support for the generalization that trainees' self-reported social accessibility relates to how well they are liked by peers and how interpersonally effective and adaptive the midboard assessment team sees the trainees to be. It is also interesting to note that self-disclosure scores were not related to the initial peer nominations but were

⁴ Scores for the three target persons were summed and only the total score used in data analysis because the individual target person scores were highly intercorrelated ($r \geq .95$ in all cases).

TABLE 1
CORRELATIONS AMONG ALL THE VARIABLES OF THE STUDY

Item	2	3	4	5	6	7	8	9	10
Self-disclosure (1)	.33*	-.34*	-.26	.06	.30*	.36**	.41**	.11	-.09
Conceptual complexity (2)		.33*	.15	.24	.23	.45**	.49**	.21	.00
Authoritarianism (3)			.00	-.02	-.10	.12	.26	.04	.19
Polarity (4)				-.24	-.17	-.26	-.14	-.30*	-.20
Peer nomination (initial) (5)					.38**	.48**	.48**	-.07	-.02
Peer nomination (6-week) (6)						.80**	.56**	.17	-.05
Interpersonal flexibility (7)							.77**	.17	-.07
General adaptability (8)								-.11	-.02
SCAT—verbal (9)									.31*
General Aptitude Test (10)									

* $p < .05$.
** $p < .01$.

related to the 6-week peer nominations. It suggests that other impression-formation type variables account for initial attitudes toward others, but that self-disclosure assumes importance in the development of more stable and less superficial interpersonal relationships. The relationship of social accessibility to the "interpersonal flexibility" and "general adaptability" ratings supports the idea that there is a process dimension underlying social accessibility which has to do with behavioral plasticity as well as openness. The correlations of conceptual complexity with both social accessibility and the adaptability ratings suggests that conceptual complexity may be tapping that dimension.

As predicted, self-disclosure was negatively correlated with authoritarianism and positively correlated with conceptual complexity, while the negative correlation with the Polarity Scale was of marginal significance ($p = .06$, two-tailed). The fact that the more authoritarian individual reports that he discloses less about himself suggests, consistent with the theory of authoritarianism, that one with relatively closed belief systems acts so as to minimize the challenge to his values and beliefs. A similar interpretation can be made of the correlation between conceptual complexity and self-disclosure. A person of a higher level of complexity presumably inter-

acts in an interdependent manner (i.e., assuming mutuality and equality in relationships; see Harvey et al., 1961), whereas one at a low level presumably interacts unilaterally (e.g., dominant and submissive roles). It follows that there should be more openness in communicating to others in an interdependent rather than a unilateral interaction.

Authoritarianism, low conceptual complexity, and the belief that human nature is evil can be viewed as three distinct cognitive-motivational bases for defensiveness towards others. It is interesting to note that together they account for about 25% of the total variance in self-disclosure (multiple $R = .51$).

While self-disclosure was related to conceptual complexity, authoritarianism, and peer nominations, the latter three variables did not correlate with each other. Thus self-disclosure is related to three aspects of interpersonal functioning which are operationally and conceptually independent, yet which all have to do with interpersonal openness and effectiveness. Those findings plus the relations of self-disclosure with the behavior ratings of interpersonal flexibility and adaptability, indicate a substantial degree of convergent construct validity for the interpretation of self-disclosure within a framework of interpersonal flexibility and openness.

Although it might be possible to interpret

some of the present results as reflecting the contribution that social desirability or other response sets make to the paper-and-pencil measures, this possibility is rather remote for several reasons: The measures of authoritarianism and polarity had specific formats designed to control for response sets, and they were not intercorrelated, which suggests they were measuring different aspects of interpersonal functioning in this study. Also response sets seem to be minimized in the Paragraph Completion Inventory because of the semi-projective nature of the responses.

The fact that neither the SCAT verbal scores nor the General Aptitude Test scores correlated with self-disclosure is one bit of evidence for the discriminant validity of self-disclosure as well.

REFERENCES

- ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality*. New York: Harper, 1950.
- CRONBACH, L. J., & MEEHL, P. E. Construct validity in psychological tests. *Psychological Bulletin*, 1955, 52, 281-302.
- FRANKFURT, L. P. The role of some individual and interpersonal factors on the acquaintance process. Unpublished doctoral dissertation, American University, 1965.
- HARVEY, O. J., HUNT, D. E., & SCHRODER, H. M. *Conceptual systems and personality organization*. New York: Wiley, 1961.
- JOURARD, S. M. Self-disclosure and other cathexis. *Journal of Abnormal and Social Psychology*, 1959, 59, 428-431.
- JOURARD, S. M. Self-disclosure and grades in nursing college. *Journal of Applied Psychology*, 1961, 45, 244-247.
- JOURARD, S. M., & LANDSMAN, M. J. Cognition, cathexis, and the "dyadic effect" in men's self-disclosing behavior. *Merrill-Palmer Quarterly*, 1960, 6, 178-186.
- JOURARD, S. M., & LASAKOW, P. Some factors in self-disclosure. *Journal of Abnormal and Social Psychology*, 1958, 56, 91-98.
- NEWCOMB, T. M. *The acquaintance process*. New York: Holt, Rinehart & Winston, 1961.
- SCHRODER, H. M., DRIVER, M. J., & STREUFERT, S. *Human information processing: Individuals and groups in complex social situations*. New York: Holt, Rinehart & Winston, 1967.
- STERN, G., CHRISTIE, R., LANE, H., SANFORD, N., & WEBSTER, H. Stereopathy-Acquiescence Scale, Form P (Personality). SSRC Stereopathy-Acquiescence Workshop, Syracuse University, 1960.
- TAYLOR, D. A. Some aspects of the development of interpersonal relationships: Social penetration processes. Unpublished doctoral dissertation, University of Delaware, 1965.
- TOMKINS, S. The psychology of knowledge. Paper presented at the meeting of the American Psychological Association, Los Angeles, September 1964.

(Received April 17, 1968)

EFFECTS OF SEX, ILLNESS, AND HOSPITALIZATION ON DAYDREAMING¹

ANN PYTKOWICZ STREISSGUTH,² NATHANIEL N. WAGNER, AND JOHN C. WECHSLER

Department of Psychiatry, University of Washington

Two hundred Ss matched for sex, in- and outpatient status, and for medical-psychiatric nonpatient status were administered a self-rating daydreaming questionnaire which was scored for frequency of daydreaming, content, and reactions to daydreams. Strong sex differences in content emerged, with males engaging in flamboyant, action-oriented daydreams, while the daydreams of women were passive, reality-oriented, and somewhat less frequent. Very few differences occurred as a function of hospitalization, but striking differences emerged on the illness dimension, with psychiatric Ss engaging in more improbable, ideational, and highly affect laden daydreams, but reacting with little concern. Although daydreaming less than psychiatric Ss, medical Ss were more frightened by and absorbed in their daydreams. In general, psychiatric Ss had a greater tolerance for emotional fantasy than did medical patients.

Experimental studies of daydreaming have dealt with a narrow range of Ss, largely middle-class college and graduate students. Although it is well documented that for this group daydreaming is a common daily experience, Singer (1966) in reviewing the daydream literature stated: "We need, therefore, to obtain some kinds of information about the true extent of inner mental activity in various samples . . . [p. 6]."

Three factors have been virtually ignored in the empirical daydream literature: sex differences,³ relationship to illness, and the effect of hospitalization. Singer (1966) reported no significant sex differences in his daydreaming studies. Pytkowicz, Wagner, and Sarason (1967) reported a sex difference that approached significance with males reporting more daydreaming. The present research investigated sex differences in both content and frequency of daydreaming and in attitudes toward daydreaming.

There are also no reported empirical studies on the relation of illness to daydreaming and little research on the effects of situational factors on daydreaming. A recent study by Antrobus, Singer, and Greenberg (1966), however, indicated that the presence of a highly distressing announcement preceding a discrimination task increased the production of spontaneous fantasy during task performance. Presumably, through fantasy, the person attempts to integrate, deal with, and adapt to the distressing information. One might postulate that similarly hospitalized patients might attempt to cope with the stress of illness through spontaneous daydreaming.

The present study, then, investigated frequency of daydreaming, content of daydreams, and attitudes toward daydreaming, as a function of sex, illness, and hospitalization. It was anticipated that men would have a greater frequency of daydreaming than women, that hospitalized patients would have a greater frequency of daydreaming than non-hospitalized patients, and that psychiatric and medical patients would have qualitatively different daydreaming patterns.

Daydreaming was defined, following Singer (1966), as any cognitive activity representing a shift of attention away from an ongoing task. Loose associations, well-developed fantasies, plans for the future, memories of the past, worries, wishes, fears, etc., were all subsumed under the term "daydreaming."

¹ This research was supported by a grant awarded to the senior author from the State of Washington Initiative 171 funds for research in Biology and Medicine.

² Requests for reprints should be sent to Ann Pytkowicz Streissguth, Child Development and Mental Retardation Center, Behavior Science Unit, University of Washington, Seattle, Washington 98105.

³ Since completion of the present study, Wagman (1967) has published a study of six differences in the content of daydreams, employing college student Ss, in which findings similar to those presented in this study are reported.

TABLE 1
CHARACTERISTICS OF SAMPLE

Group	Age in years			Education		
	20-34	35-49	50+	-11	12-14	15+
Psychiatric inpatient						
Male	9	10	1	2	9	9
Female	9	9	2	2	11	7
Psychiatric outpatient						
Male	12	7	1	3	8	9
Female	9	9	2	3	9	8
Medical inpatient						
Male	10	9	1	3	10	7
Female	9	10	1	2	10	8
Medical outpatient						
Male	8	10	2	2	8	10
Female	8	11	1	2	9	9
Nonpatients						
Male	9	10	1	3	10	7
Female	9	10	1	2	10	8
Total	92	95	13	24	94	82

Note.—*N* = 20 for each group; total *N* = 200.

METHOD

Subjects

The *Ss* were selected from three Seattle, Washington, hospitals, and were equally divided within the following subgroups: psychiatric and medical patients and controls (nonpatients), in- and outpatients, and men and women. The psychiatric *Ss* were obtained first and the medical *Ss* and controls were then matched according to age and education. A purposefully wide range of *Ss* were selected, with age ranging from 20 to 73 years with a mean of 35.9 years and education ranging from 7 to 23 years with a mean of 14.1 years.

Certain unavoidable differences occurred between subgroups that reflect actual differences between the services from which *Ss* were selected.

Length of hospitalization varied for each subgroup according to treatment procedures on each service. Psychiatric inpatients ranged in length of hospitalization from 3 to 120 days with a mean of 33.83 days. Medical inpatients' length of hospitalization ranged from 1 to 48 days with a mean of 5.05 days.

Diagnoses of the psychiatric *Ss* included: 35% schizophrenics, 32% neurotic reactions, 24% character disorders, and 9% affective psychoses. There were somewhat more schizophrenics in the inpatient group and more neurotic reactions in the outpatient group, reflecting the nature of inpatient psychiatric patients compared to those seen in outpatient clinics. The diagnoses of medical *Ss* reflected the services from which they were obtained: 23% cardiology, 19% gastrointestinal, 12% arthritis, 11% chest and infectious diseases, 9% metabolic and endocrine, 6% renal, 5% dermatology, 5% hematology, 5% allergy, and 5% general medical. When tested, the control *Ss*

were not patients and had no history of psychiatric treatment. They were recruited predominantly from hospital personnel and, where necessary, from acquaintances of the *Es*.

The sample as a whole had the following marital, ethnic, and religious characteristics which do not appear to differ between subgroups: 65% were married, 35% were single, divorced, or widowed; 29% were from British backgrounds, 28% from Western European, 16% from Scandinavian, 6% from Eastern European, and 21% from American or "other"; 68% were Protestant, 13% Catholic, 4% Jewish, and 15% "other" or had no religious preference. The sample is remarkably representative of the Seattle area.

Procedure

Each *S* was given the Singer Daydream Questionnaire,⁴ which contains 120 specific daydreams to which an *S* responds on a 1-5 scale according to the frequency with which he has a particular dream.⁵ In addition, there are 142 statements about daydreams to which *Ss* respond on a 5-point scale ranging from emphatically true to emphatically false.

The *Ss* were given the test either individually or in small groups. In each case, the study was explained as a research project concerning what people daydream and how they feel about their daydreams. The

⁴ Now revised and published in a new form entitled "Imaginal Processes Inventory" by Singer and Antrobus, published by the City College of New York.

⁵ Although this study used the form of the Daydream Questionnaire used by Singer in his early studies, the directions were those described by Pytkowicz et al., 1967.

examiner was available for questioning, but answered questions in a nonspecific but supportive manner.

Experimental Design

The Ss formed a $2 \times 2 \times 3$ factorial design with the following independent variables: sex (male vs. female); hospitalization (outpatients vs. inpatients); and illness (psychiatric vs. medical patients vs. controls). Since there were, of course, no inpatient controls, this resulted in 10 subgroups of 20 Ss each. Scores were derived from the Singer Daydream Questionnaire. Thirty dependent variables were studied, including one scale of daydreaming frequency, 11 scales of daydream content, and 18 scales of attitudes about and attributes of daydreams.

The necessary absence of the inpatient controls meant that the usual three-way analysis of variance had to be modified to allow for the unbalanced design (Dixon, 1965, p. 543; Scheffé, 1963). In particular, it was necessary to adjust the means of the control group to make them comparable with the psychiatric and medical groups which included inpatients. The adjusted control mean is given by:

$$\bar{y}_{\text{control}} + \frac{\bar{y}_{\text{in}} - \bar{y}_{\text{out}}}{2}$$

where \bar{y}_{control} is the mean for all control Ss while \bar{y}_{in} and \bar{y}_{out} are the means for inpatients and outpatients, respectively, in the psychiatric and medical groups. The comparison of inpatients and outpatients involved only the use of psychiatric and medical Ss while the sex comparison used all Ss.

TABLE 2

ANALYSIS OF VARIANCE OF MALE VERSUS FEMALE DIFFERENCES ON THE DAYDREAMING SCALES

Variable ^a (Daydream scale)	Mean scores		F
	Male	Female	
Sexual	17.48	9.60	18.58**
Heroic	6.30	3.40	14.92**
Modified aggressive	6.62	4.55	10.94**
Acceptance of daydreams	31.48	25.92	10.70**
Repetitiveness of daydreams	15.90	12.59	9.19**
Achievement-oriented	8.27	6.52	8.83**
Factor H	55.30	47.24	8.79**
Probable	10.33	12.85	8.28**
General emotional response	16.32	14.13	6.76**
Present-orientation	14.92	12.77	5.98*
Frequency	107.67	90.52	3.74
Passivity	21.06	22.86	3.69
Conscience	5.05	4.03	3.66
Problem solving	8.37	6.76	3.38
Attempt to control	19.47	17.34	2.20

Note.—N = 20.

^a Arranged in order of magnitude of difference.

* $p < .05$.

** $p < .01$.

Psychiatric, medical, and adjusted control means were compared by Duncan's multiple-range test with special modifications for the unbalanced design (Duncan, 1955). Because the adjusted mean for the control group had a larger variance than the psychiatric and medical means, comparisons involving the control mean were less likely to reach statistical significance.⁶

RESULTS

Differences as a Function of Sex

Overall sex differences in frequency of daydreaming did not attain statistical significance as shown in Table 2. The trend of the sex difference in frequency replicates a previous finding (Pytkowicz et al., 1967) in which male college students also reported somewhat more daydreaming than women. The Ss in this study, spanning a much wider age and diagnostic range, have a lower overall frequency of daydreaming compared with those in the earlier study, although the male-female relationship remains the same.

The content of men's daydreams differ significantly from the daydreams of women along a dimension of activity-passivity. Daydream scales endorsed significantly more frequently by men include sexual daydreams, heroic daydreams, achievement-oriented daydreams, and modified aggressive daydreams. The only content scale endorsed significantly more often by women than men is probable daydreams, which includes a selection of rather concrete, practical daydreams such as planning what to cook for dinner or what clothes to wear. Women also tend to score higher than men on a scale of passivity in daydreaming, which includes a selection of fairly static, unimaginative, inactive daydreams (this latter difference, however, does not achieve significance at the .05 level).

In attitudes toward daydreaming, men score significantly higher than women on the following scales: acceptance of daydreaming, repetitiveness of daydreams, general emotional response to daydreaming, and present-orientation in daydreams. Scales on which

⁶ The authors are most grateful to Ethel Schaeffer, Department of Preventive Medicine, University of Washington, for her assistance with the statistical analyses. They also wish to acknowledge the assistance of James Oakland and Marjorie Rubin in the collection of the data.

there is a trend for men to endorse more than women are: scales of conscience daydreams and problem solving through daydreaming. Many items on the latter scale are suggestive of daydreams being used actively in the creative process (such as to find new ways of approaching a problem).

A noteworthy finding in terms of sex differences in daydreaming is the absence of any differences on such scales as anxious daydreams, improbable daydreams, overt-aggressive daydreams, pregenital daydreams, and need-satisfaction daydreams.

Men had higher scores than women on Factor H, a factor with substantial loadings on neuroticism and other pathological non-daydreaming scales (Singer & Antrobus, 1963). There is no reason to assume higher scores for males reflect increased neuroticism, raising questions, therefore, about the interpretation of this factor.

In summary, the daydreams of men differ along a number of dimensions from those reported by women. The primary difference occurs in terms of the action-oriented and imaginative content of men's daydreams compared to the passive, reality-oriented daydreams of women. Men also are more accepting of their daydreams, more emotionally involved and appear to use daydreaming for new approaches to problem solving.

Although it is possible that the male-female differences reported here are primarily a function of the patient status of the majority of Ss in this study, the lack of any significant interaction effects contraindicates this interpretation. Further, in a study (Wagman, 1967) which appeared after this study was completed, sex differences of college students were strikingly consistent with the findings of our sample.

Differences as a Function of Hospitalization

This study reveals few differences in the daydreaming patterns of hospitalized compared with nonhospitalized patients (see Table 3). Hospitalized patients are significantly more accepting of their daydreams and report significantly more problem-solving daydreams. On the other hand, hospitalized patients report significantly fewer sexual and angry-emotional daydreams than do outpa-

TABLE 3

ANALYSIS OF VARIANCE OF INPATIENT VERSUS
OUTPATIENT DIFFERENCES ON THE
DAYDREAMING SCALES

Variable ^a (Daydream scale)	Mean scores		F
	Inpatients	Outpatients	
Acceptance of daydreams	32.05	26.68	13.00**
Sexual	11.35	15.93	7.86**
Angry-emotional	6.34	8.54	6.91**
Problem solving	8.83	7.35	4.46*
Frequency	91.15	107.74	3.62
Physical sensation in daydreaming	22.25	20.60	2.80
Visual imagery	9.24	8.09	2.68
Anxious	6.35	7.46	2.40

Note.—N = 200.

^a Arranged in order of magnitude of difference.

* $p < .05$.

** $p < .01$.

tients. No significant differences in frequency of daydreaming occurred, although the trend for hospitalized patients to daydream less than nonhospitalized patients was surprising.

To explore this further, each subgroup of hospitalized patients, medical and psychiatric, were rank-ordered separately according to length of hospitalization at time of testing. Each subgroup was divided by a median split into long and short duration of hospitalization. Male and female Ss were distributed evenly throughout long and short subgroups for each diagnostic category, thus ruling out possible interaction effects between sex, length of hospitalization, and frequency of daydreaming.

TABLE 4

FREQUENCY OF DAYDREAMING AS A FUNCTION
OF LENGTH OF HOSPITALIZATION

Item	Length of hospitalization			
	Medical inpatients		Psychiatric inpatients	
	Short	Long	Short	Long
Mean number of days in hospital	2.1	9.8	11.8	51.9
Mean score: frequency of daydreaming	69.5	98.3	114.1	95.4

As Table 4 indicates, length of hospitalization has clearly opposite effects on frequency of daydreaming for psychiatric compared with medical patients. Medical patients daydream more the longer they are hospitalized while psychiatric patients start out with a higher level of daydreaming, and daydream less as length of hospitalization increases. These findings are obviously confounded by the lack of equivalence between the short and long categories of hospitalization for the medical compared with the psychiatric patients. It cannot be ascertained from these data whether hospitalization has a differential effect on daydreaming according to the type of illness manifested by the patient, or whether length of hospitalization has a curvilinear effect on daydreaming frequency irrespective of type of illness. Research controlling for length of hospitalization seems warranted to clarify this point.

Differences as a Function of Illness

By far the most striking differences in daydreaming patterns occur as a function of type of illness. Table 5 shows the analysis of vari-

ance of the three categories of illness (psychiatric, medical, and controls) for the daydreaming scales on which near-significant overall differences were obtained. Due to the difficulty in interpreting this three-way analysis, a Duncan's multiple-range test, with special modifications as described above, was employed.

In terms of frequency of daydreaming, psychiatric patients clearly report more daydreaming than either controls or medical patients. Although medical patients report less daydreaming than controls, this latter difference is not significant.

A large cluster of scales having a highly affective content are endorsed most frequently by psychiatric patients and least frequently by medical patients. These include: improbable daydreams, anxious daydreams, conscience daydreams, sexual daydreams, overt-aggressive daydreams, modified-aggressive daydreams, and angry-emotional daydreams. This cluster might be regarded as daydreams most highly characteristic of the psychiatric patients and least characteristic of medical patients.

The only scale characteristic of medical patients only (the scale on which medical patients score highest and psychiatric patients score lowest) is absorption in daydreaming. Medical patients also score significantly higher than psychiatric patients on the following scales: Factor H, frightened reaction to daydreams, and repetitiveness of daydreams.

It is clear that psychiatric and medical patients not only engage in different kinds of daydreams, but they have markedly different reactions to the experience of daydreaming. Psychiatric patients not only daydream more frequently than medical patients and nonpatients, but their daydreams are more improbable (including more fantastic and flamboyant themes and an existence in impossible roles), more conscience-ridden (including themes of punishment and censure for personal wrongdoing), and include more themes of a pre-genital and sexual nature (including narcissistic, self-aggrandizing, and self-gratifying themes such as owning expensive clothing, eating at great banquets, and being handsome and irresistible). Psychiatric patients also have

TABLE 5

ANALYSIS OF VARIANCE OF PSYCHIATRIC PATIENTS
VERSUS MEDICAL PATIENTS VERSUS
CONTROLS

Variable (Daydreaming scale)	Mean score *			F
	Psychiatric	Control	Medical	
Anxious	8.63	6.40	5.19	12.19**
Angry-emotional	9.43	7.52	5.45	11.76**
Modified aggressive	7.12	5.35	4.16	10.57**
Overt aggressive	5.50	4.10	3.00	10.38**
Frequency	117.99	97.70	80.90	9.50**
Sexual	16.89	13.15	10.39	8.90**
Conscience	6.01	4.08	3.30	8.77**
Frightened reaction	7.44	11.57	12.48	8.05**
Absorption in daydreams	17.49	18.10	21.19	7.10**
Improbable	5.76	4.08	3.79	7.10**
Pregenital	5.43	3.00	3.11	6.91**
Achievement-oriented	8.50	7.98	6.00	6.42**
Heroic	6.08	5.30	3.40	6.39**
Repetitiveness	13.01	13.50	15.85	3.96*
Future oriented	18.75	15.05	18.36	3.60*
Factor H	48.14	48.85	52.06	3.51*

Note.—N = 200.

* Means connected with a single line are not significantly different from each other at the $p < .05$ level according to Duncan's multiple-range test.

** $p < .05$.

*** $p < .01$.

more highly affective and action-oriented daydreams, including overtly aggressive disfigurement daydreams as well as more modified aggressive or nonphysically aggressive daydreams and more angry-emotional daydreams. They also have more anxious daydreams, more heroic, and more achievement-oriented daydreams. Despite the highly affect-laden content of the daydreams of psychiatric patients, they are less frightened by their daydreams, less absorbed in daydreaming, and experience less repetitiveness of daydreaming.

By contrast, medical patients (who do *not* differ from nonpatients in terms of overall frequency of daydreaming, nor in the amount of improbable, sexual, or pregenital daydreams) *do* report significantly less of the highly affect-laden daydreams than even the controls. Medical patients admit to significantly less aggressive, anxious, heroic, angry-emotional, and achievement-oriented daydreams than either psychiatric patients or controls.

Despite the relative absence of themes of a highly ideational or affective nature in the daydreams of medical patients and the smaller frequency of daydreaming in medical patients, it is notable that medical patients are apparently more bothered by their daydreaming activity. They have a greater emotional response to daydreaming, are more absorbed in their daydreams, experience the most repetitive daydreams, and are most frightened by their daydreams.

In summary, the reported fantasy life of psychiatric patients is characterized by a strong ideational component in which specific, highly affective, and fantastic themes are dominant. Although the nature of their daydreams is highly affect-laden, psychiatric patients report little emotional reaction to daydreaming. Medical patients, on the other hand, experience little specific ideational or affective content in their daydreams, but do report a strong emotional reaction to daydreaming including being more absorbed in and frightened by their fantasy life.

DISCUSSION

This study supports a concept of daydreaming as a personality characteristic on which a wide variety of human experiences

can be played. Sex differences in reported patterns of daydreaming seem to reflect quite directly the cultural stereotypes of male-female interests in the United States. Men engage in action-oriented, imaginative fantasies modeling the cultural role expectations of the male, while the women engage in probable, reality-oriented, and passive daydreams, reflecting the accepted feminine roles of passivity, attention to detail, and practicality. Thus, the respective daydreams of adult men and women appear as replications of the sex-role play of young children where the girls play house and have tea parties while the boys play pirates and take rocketships to the moon.

Two questions should be raised regarding any such study of daydreaming as the present one: first, to what extent the reported daydreams of Ss are representative of their actual fantasy life, and to what extent their reported daydreams may be related to their perceptions of what is socially desirable. There is little point in belaboring the first point, as it is difficult to imagine any way other than self-reports to measure an individual's daydreaming, so all that would ever be measurable would be reported daydreaming activity. Regarding the second question, both Singer (1966) and Pytkowicz et al. (1967), report no significant relationship between the frequency of daydreaming and the construct of social desirability.

The hospitalization effects are particularly interesting in view of other studies dealing with the relationship of decreased sensory stimulation to imagination. Studies of the fantasy life of blind children, for example, indicate that although they do engage in daydreams, night dreams, and internal monologue, the entire fabric of their fantasies is less rich and imaginative and more concrete, simple, and reality-bound (Singer & Streiner, 1966). Similarly, children raised in institutions have a notable paucity of imaginative resources (Goldfarb, 1949). Singer (1966) points to the necessity for both a richness of external stimulation and an opportunity for privacy in the development of an active fantasy life. In this respect, the tendency for hospitalized patients to not only daydream significantly less about highly affect-laden sexual and aggressive themes as well as to be more accept-

ing of daydreams and to use daydreaming for problem solving is interesting.

The anticipation that hospitalized patients might use daydreaming as an effective device for coping with the stress of hospitalization is not supported by these data. Procedural factors may be more at fault than the proposition itself. It is quite possible that the Singer Daydream Scale is not sensitive to short-term changes in daydreaming activity and may be measuring more enduring personality characteristics. It is also possible, indeed suggested by this study, that length of hospitalization may be a more salient factor than the experience of hospitalization per se.

This is further complicated by the fact that psychiatric hospitalization is almost invariably longer than medical hospitalization due to the different goals of treatment during the hospitalization. Fantasy as a mechanism for coping with illness may well have been activated at the outset of hospitalization or at the time of diagnosis and planning. Certain characteristics of the hospital system which encourage the relinquishment of responsibility and self-control on the part of the patient may actually inhibit spontaneous fantasy. Daydreaming may well be a manifestation of an individual's feeling of personal choice and responsibility in a given situation—an attempt to work through possible alternatives and ideas, rather than an aimless flight of ideas in an apathetic state. Differential use of daydreaming as a function of length of hospitalization in medical compared with psychiatric patients is suggested by this study. Length of hospitalization, severity of illness, and individual differences of Ss are all factors needing further investigation to understand the effect of hospitalization on spontaneous fantasy. The restitutive and protective functions of fantasy do not appear to have been adequately measured in the present study.

The striking differences between the daydreaming patterns of medical and psychiatric patients as well as their marked differences in reaction to the experience of daydreaming are provocative. While obviously no cause and effect interpretations can result from these data, it is interesting to speculate to what extent these distinctive daydream patterns are characteristic of the cognitive life styles of patients

who develop psychiatric symptoms compared with those whose symptoms are medical. Not only do the psychiatric patients apparently involve many of their worries and conflicts in their active fantasy life, but they do so with relatively little concern. Anxiety and affect is focused in the theme of the daydream rather than the experience of daydreaming.

By contrast, medical patients deny active imaginative fantasies. Their daydreams, rather than being highly ideational, are loosely organized, ruminative, repetitive, and egocentric. While denying affectively-toned themes, the medical patients report much reactivity and concern over the process of daydreaming, apparently attempting to consciously control and diminish this distasteful and frightening activity. The repressive and suppressive quality that pervades their fantasy life may reflect the fear of individuals facing crucial health problems, or it may reflect a style of life that results in physical illness as a response to stress (Grace & Graham, 1952; Hinkle & Wolff, 1957).

The view that daydreaming is in itself pathological is not supported by these data. Consistently, the mean scores of nonpatient control Ss fall between the mean scores of the psychiatric patients and the mean scores of the medical patients. If the spontaneous fantasy of the control Ss are taken as a norm, then the psychiatric and medical patients repeatedly, on almost all of the 30 scales, deviate in opposite directions from the norm. Although psychiatric patients report more daydreaming, there is no evidence from this study to suggest that the process of daydreaming is uniquely associated with mental illness. Rather it appears more meaningful to conceptualize daydreaming as a cognitive activity, the quantity and quality of which is associated with a variety of personal dimensions of which two are sex and type of illness.

REFERENCES

- ANTROBUS, J. S., SINGER, J. L., & GREENBERG, S. Studies in the stream of consciousness: Experimental enhancement and suppression of spontaneous cognitive processes. *Perceptual and Motor Skills*, 1966, 23, 399-417.
- DIXON, W. J. (Ed.) *Biomedical computer programs*. Health Sciences Computing Facility, Department of

- Preventive Medicine and Public Health, School of Medicine, University of California, Los Angeles, California, 1965.
- DUNCAN, D. B. Multiple range tests for correlated and heteroscedastic means. *Biometrics*, 1955, 11, 1-42.
- GOLDFARB, W. Rorschach test differences between family-reared, institution-reared, and schizophrenic children. *American Journal of Orthopsychiatry*, 1949, 19, 624-633.
- GRACE, W. J., & GRAHAM, D. T. Relationship of specific attitudes and emotions to certain bodily diseases. *Psychosomatic Medicine*, 1952, 14, 243-251.
- HINKLE, L. E., & WOLFF, H. G. The nature of man's adaptation to this total environment and the relation of this to illness. *AMA Archives of Internal Medicine*, 1957, 99, 442-460.
- PYTROWICZ, A. R., WAGNER, N. N., & SARASON, I. G. An experimental study of the reduction of hostility through fantasy. *Journal of Personality and Social Psychology*, 1967, 5, 295-303.
- SCHEFFÉ, H. *The analysis of variance*. New York: Wiley, 1963.
- SINGER, J. L. *Daydreaming: An introduction to the experimental study of inner experience*. New York: Random House, 1966.
- SINGER, J. L., & ANTROBUS, J. S. A factor-analytic study of daydreaming and conceptually-related cognitive and personality variables. *Perceptual and Motor Skills*, 1963, 17, 187-209.
- SINGER, J. L., & STREINER, B. Imaginative content in the dream and fantasy play of blind and sighted children. *Perceptual and Motor Skills*, 1966, 22, 475-482.
- WAGMAN, M. Sex differences in types of daydreams. *Journal of Personality and Social Psychology*, 1967, 7, 329-332.

(Received April 29, 1968)

Manuscripts Accepted for Publication in the

Journal of Consulting and Clinical Psychology

This listing plus those published in preceding issues represents the backlog of manuscripts accepted by this Journal. Such listing is intended to allow readers to become aware of research many months in advance of journal publication.

- Driver Accidents and the Neuropsychiatric Patient: Matthew W. Buttiglieri,* M. I. Charles E. Woodson, Marie Guenette, and Mae Thomson: Veterans Administration Hospital, Sepulveda, California 91343.
- Perceived Child-Rearing Practices and Self-Disclosure Patterns: Joseph R. Doster and Bonnie Ruth Strickland*: Department of Psychology, Emory University, Atlanta, Georgia 30322.
- Sensing Humor: Latency and Amplitude of Response Related to MMPI Profiles: Douglas S. Holmes*: Smith Richardson Foundation, Inc., P.O. Box 3265, Greensboro, North Carolina 27402.
- Patterns of Performance on WISC Similarities in Emotionally Disturbed and Brain-Damaged Children: Leon P. Hall* and M. LaVerne Driere: Oakland Schools, Campus Drive-County Service Center, Pontiac, Michigan 48053.
- The Psychological Characteristics of Neighborhood Youth Corps Enrollees: Sherman Eisenthal* and Lewis J. Sherman: Boston University School of Medicine, 80 East Concord Street, Boston, Massachusetts 02118.
- The Sensation-Seeking Scale as a Predictor of Need for Stimulation during Sensory Restriction: John Hocking and Malcolm Robertson*: Psychology Department, Western Michigan University, Kalamazoo, Michigan 49001.
- Intelligence of Volunteers as Research Subjects: Jay M. Cudrin*: Department of Psychology, Connecticut Valley Hospital, Middletown, Connecticut 06457.
- Humor and Suicide: Favorite Jokes of Suicidal Patients: Don Spiegel,* Patricia Keith-Spiegel, Joel Abrahams, and Lionel Kranitz: Brentwood Hospital (183B), Veterans Administration Center, Los Angeles, California 90073.

* Asterisk indicates author for whom address is supplied.

USE OF SPEECH IN THE INHIBITORY CONTROL OF MOTORIC BEHAVIOR OF SCHIZOPHRENIC PATIENTS¹

BEATRICE A. ASHEM²

University of Waterloo

The conditions under which speech influences the speed of a tapping response in schizophrenics ($N = 48$) and nonschizophrenics ($N = 48$) were experimentally investigated. The Ss verbalized three words ("goads") under varied pretask instructions (Minimal and Slow Instructions) and two sources of delivery (self-delivered "Internal" verbalizations and other-delivered "External" verbalizations). The results indicated that all groups inhibited speed of tapping for a Slow Instruction. The schizophrenics under the Internal goad condition inhibited response as well as the control group; under the External goad condition these patients were less able to inhibit response. For the Minimal Instruction condition the schizophrenics responded to the meaning inherent in the goads "faster" and "slower" for the Internal but not the External goad condition. The converse was true for the control group.

Several papers in the recent literature have been concerned with the relationship between the slow performance which characterizes schizophrenic patients on simple motor tasks and situationally relevant variables. The search for relevant dimensions influencing performance separates into two categories on the basis of (a) identification of stimulus components and (b) those concerned with measures of response-mediated feedback.

The experiments in the first category have indicated that deficit performance of schizophrenics on a reaction-time (RT) task can be eliminated by increasing the intensity of

the discriminative stimulus (Wienckowski, 1959), or reducing distracting elements of the stimulus configuration (Payne & Caird, 1967); those of the second category have found improvement on an RT task to be associated with increased information relative to the adequacy of the response (Lang, 1959). Buss and Lang (1967), following an extensive review of the literature, have suggested that if the schizophrenic is given more information relevant to his performance than is usually required by normal Ss his performance is adequate. This holds true for verbal conditioning and cognitive learning as well as motor performance. The authors hypothesize that the greater reliance of the schizophrenic upon external cues may be due to a breakdown in self-instructional set.

The import of isolating the variables influencing performance of the schizophrenic on any task is clear. The problem remaining to plague the investigators is whether the improved performance gained in the laboratory can be maintained outside of the experimental setting.

Luria (1966) maintains that self-control (or internalized control) required for appropriate responding to varying environmental demands in humans is attained only following the development of speech mechanisms. One might reasonably ask whether self-speech might be used to organize and control motor behaviors of schizophrenics. Two studies (Birch, 1966; Lovaas, 1964) dealing with

¹ This paper is based, in part, on a dissertation submitted to the University of Waterloo in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author wishes to express her gratitude to her research advisor, R. A. Steffy, and other committee members, D. Meichenbaum, B. Gilmore, and K. Bowers for their many helpful suggestions. Special thanks are extended to C. M. Franks for his encouragement and support in data collection at the New Jersey Neuropsychiatric Institute, Princeton, New Jersey, and to the Medical Directors and Psychology Departments of the following state hospitals in New Jersey: Greystone State Hospital, Marlboro State Hospital, and Trenton State Hospital. Thanks are extended as well to J. A. Rosenkrantz, Newark Beth Israel Hospital, Newark, New Jersey, for permission to collect data for a control group. A portion of the support of this research came from the Ontario Mental Health Foundation, Project No. 76.

² Requests for reprints should be sent to Beatrice A. Ashem, now at the Behavior Therapy Unit, Douglas Hospital, 6875 LaSalle Boulevard, Montreal, P.Q., Canada.

the symbolic representation of speech to influence the motor behavior of children are relevant here. Both demonstrate that self-speech plays a central role in the inhibition and facilitation of simple motor responses.

Lovaas (1964) had children bar-press while saying the words "faster" and "slower." It was shown that there was a relationship between the facilitatory and inhibitory properties of the words and developmental level; the older children responded to the meaning component of the words, while younger children did not. It should be noted that pretraining was used, with reward contingent upon counting fast or slow in the presence of a light, to establish a "meaning" effect to the bar-press response.

The purpose of the present study was to examine the conditions under which speech controls the speed of a simple continuous tapping response in schizophrenics and non-schizophrenics. The experiment differs from other studies attempting to delineate the variables influencing motor performance of schizophrenics because it uses as a dimension of study *S*'s own speech as well as that of others. Further it examines the effects on the response of a pretask instruction which is consistent, irrelevant, or contradictory to words ("goads") spoken during actual performance of a task. A preliminary investigation suggested that the meanings of the words "faster" and "slower" influenced the tapping performance of schizophrenics and normals when *Ss* themselves spoke these words (Internal condition). When *E* spoke the words (External condition), normals modified performance according to the meaning of the words but the schizophrenics did not respond to the meaning component of the words. Further, performance appeared to be modified to an important extent by a pretask instruction. For a Fast Instruction condition (e.g., *Ss* were asked to tap as fast as they could while verbalizations were occurring) no differences between the schizophrenics and normals were obtained; for a Slow Instruction condition, however, schizophrenics were significantly less able than normals to tap slowly when *E* spoke the goads (External condition) even though these patients were able to do so when they spoke the goads themselves.

The present study is a controlled examination of the parameters involved in verbal control over a simple motor response in schizophrenics and nonschizophrenics. As the Fast Instruction condition had not yielded significant differences between schizophrenics and normals in the preliminary investigation, this condition was not repeated.

METHOD

Subjects

The 48 *Ss* of the schizophrenic (*S*) group were selected from admission wards of four state hospitals in New Jersey according to the following criteria: a diagnosis of schizophrenia with no history of complicating neurological or physical disorders, male, between the ages of 23 and 45 years, length of hospitalization not over 3 months, no more than three previous hospitalizations, any previous hospitalizations not to exceed 3 months, not about to be discharged, no EST therapy within a 10-week period, nonassaultive and able to cooperate. All patients were on phenothiazine drug treatment at the time of testing. Dosages were noted, and no patient with unusually high dosages was included. The mean age was 30.8. The mean IQ derived from a WAIS prorated vocabulary subtest was 105.4.

The 48 *Ss* of the nonschizophrenic (*C*) group were 21 convalescent patients at the Newark Beth Israel Hospital, 19 alcoholics who were voluntary admissions to the New Jersey State Hospitals, and 8 hospital attendants. The criteria for selection of *Ss* were: hospitalized at time of testing, between the ages of 23 and 45 years, male, no known history of psychiatric disorders, able to sit up freely, in no physical pain at the time of testing. Due to the difficulty of obtaining hospitalized patients to satisfy the established criteria eight hospital attendants were added to complete the sample. The mean age was 34.0; the mean IQ, 103.8.

There was no danger of a possible response bias resulting from communication to one another of the nature of the task, as general hospital patients were from widely separated wards. To avoid such bias among the remainder of the *Ss* each was separated from the untested portion of the group as soon as testing was completed.

Selection Tests

All *Ss* were administered the Vocabulary subtest of the Wechsler Adult Intelligence Scale (WAIS); the Ullmann-Giovanni Self-Report Scale (Ullmann & Giovanni, 1964) and the Zigler Social Competence Scale (Zigler & Phillips, 1962) just prior to the experimental test. The *Ss* within the two major *S* groups were then matched on both the Ullmann-Giovanni Self-Report Scale (Ullmann) and the Zigler Social Competence Scale (Zigler). The purpose of the test was to ensure that differences obtained between groups for the different experimental condi-

tions could not be attributed to bias deriving from individual differences in intellectual and social abilities. Separate analyses of variance of scores obtained on the Zigler and on the Ullmann showed the S group scored lower than the C group on both scales ($F = 26$, $df = 1/88$ and $F = 24.6$, $df = 1/88$, respectively). The differences between the groups were significant beyond the .001 level. The differences in age and IQ were not significant.

Apparatus

The equipment consisted of a tapping board, an attached power supply and automatic counter, and a tape recorder. The tapping board had a wooden base, $9\frac{1}{2}$ inches \times $9\frac{1}{2}$ inches, to which was attached a metallic plate, 5 inches square. A pencil, with a dual electric cord to the metallic plate and to the power supply, was attached. When the pencil touched the metallic plate it automatically activated the counter.

The experiment took place in an office on the wards where the individual patients were located.

Experimental Design

There were three independent variables: level of pathology, source of goad, and pretask instruction. The level of pathology variable was divided into two conditions: schizophrenics and nonschizophrenics; the source of goad variable was divided into the Internal and External conditions; the pretask instruction variable was divided into the Minimal Instruction and Slow Instruction conditions. For each condition, Ss faced verbal commands (goads) which were of three levels of consistency with the pretask instruction. One goad condition (the word "slower") had high consistency with the Slow Instruction condition; the second condition (the word "letter") was irrelevant; the third ("faster"), contradictory. A fourth condition within each group was a control condition using no word (instructional set only). There was a total of eight S groups with 12 Ss in each group. Assignment of Ss to the experimental conditions was random.

The goads were used by all Ss within each treatment condition in counterbalanced order. One-half of the Ss in each experimental condition was given the goads in the order "letter-faster-slower-letter"; the remaining half, in the order "letter-slower-faster-letter." The "no-word" condition was used to complete each series, that is, the presentation of the within-conditions was in the order "letter-(faster)-(slower)-letter-no word."

The dependent variable was the number of taps averaged over two 30-second intervals.

In summary, the Ss were placed into a six repeated measurements (including the operant level of tapping), $2 \times 2 \times 2$ between-within analysis of variance design.

The number of times the verbal goad was presented in a 30-second tapping trial was, in the Internal condition, determined by the Ss themselves. For the External condition the number of times the

verbal goad was presented was determined by yoking each S with a "partner" in the Internal condition.

The goads "letter," "faster," and "slower" were rated on a 7-point rating scale by each S at the end of the experimental session to evaluate the speed connotations which each of the goads possessed for the two S groups.

Pretraining

To ensure that Ss would understand that the goads to be used actually had speed connotations, all Ss received pretraining on the goads.

The Ss were asked to say a word over and over for 15 seconds. The goad "letter" was given first. Four trials were given, and the number of repetitions of the goad were recorded by E. "Faster" or "slower" in counterbalanced order was given next. The criterion established for "faster" was a minimum of 24 repetitions per 15-second interval; for "slower" the criterion was a maximum of eight repetitions per 15-second interval. The S was told to say the word according to the meaning of the word.

Procedure

At the beginning of the experimental session, Ss were told that they were participating in an investigation of finger movements and muscular coordination. The equipment was pointed out to them and briefly explained.

The S was then told "I would like you to tap until I tell you to stop. Be sure to raise the pen from the board after each tap, and move across the board this way. Any questions? Ready? Begin." Two trials were obtained, which served as the operant level of responding. The treatment to which S had been assigned was then applied.

Minimal Instruction—Internal goad condition: Each S was told, "Now I would like you to tap while you are saying a word. Say the word all the while you are tapping—the word is 'letter' (S was shown the written word, and asked to enunciate it). Say the word 'letter' aloud all the while you are tapping. Ready? Begin." The S was then told, "Now I would like you to say another word. The word is 'faster' ('slower'). You know what the word 'faster' ('slower') means. I am going to ask you to tap as you are saying the word 'faster' ('slower'). Ready? Begin." These instructions were repeated for the second word, "slower" (or "faster"), whichever occurred in the counterbalanced order.

On the next trial Ss were asked to tap while saying the word "letter" again. This trial was included to assess any effects due to fatigue or practice. On the final trial, "no word" condition, Ss were asked to tap without saying anything.

Minimal Instructions—External goad condition: The procedure and instructions were identical except for the appropriate changes. For instance, "Now I would like you to tap while I am saying a word. I am going to say the word all the while you are tapping—the word is 'letter.' I am going to say the word 'letter' aloud as you tap. Ready? Begin."

TABLE 1
MEAN TRIALS TO CRITERION ON PRETRAINING

Group	Minimal instructions				Slow instructions			
	Internal		External		Internal		External	
	Faster	Slower	Faster	Slower	Faster	Slower	Faster	Slower
Schizophrenics	1.66	3.90	1.59	5.09	2.51	4.09	2.52	5.09
Nonschizophrenics	1.84	4.50	1.53	5.50	2.09	4.51	1.61	3.73

Slow Instructions—Internal goad condition: Each S was told, "I would like you to tap as slowly as you can while you are saying a word. Say the word all the while you are tapping—the word is 'letter' (demonstration and enunciation). Say the word 'letter' aloud as you tap. Now tap as slowly as you can while you are saying the word 'letter.' Ready? Begin." Then S was told, "Now I would like you to tap as slowly as you can while you are saying another word. The word is 'slower' ('faster'). You know what the word really means. Say the word 'slower' ('faster') aloud as you tap. I want you to tap as slowly as you can while you are saying 'slower' ('faster'). Ready? Begin."

On the next trial Ss were asked to tap as slowly as they could while saying the word "letter." On the final trial, "no word" condition, Ss were asked to tap as slowly as they could without saying anything.

Slow Instructions—External goad condition: The procedure and instructions were identical to those under the Slow Instruction—Internal goad condition, except for the appropriate changes as described under Minimal Instructions, External condition.

The goads spoken by E were pronounced in the same manner throughout the experiment. A monotonous tone was used rather than an expressive one which might have conveyed a meaning to S.

RESULTS

Pretraining Data

Table 1 presents the mean trials to criterion required by the schizophrenic (S) and control (C) groups on pretraining in the pacing of the words "faster" and "slower." As there were no order effects from counterbalancing the words for either of the groups, the data were combined for further analysis. Separate analyses of the trials required to criterion for the words "faster" and "slower" revealed no differences between the S and C groups ($F = 0.09$, $df = 1/88$ and $F = 0.9$, $df = 1/88$, respectively). Both groups required twice as

TABLE 2
MEANS AND STANDARD DEVIATIONS OF TAPPING SCORES FOR ALL EXPERIMENTAL CONDITIONS

Instructions	Internal						External					
	O	L	F	S	L	N	O	L	F	S	L	N
Minimal												
Schizophrenics												
<i>M</i>	115.1	68.0	81.7	63.5	72.6	125.7	125.4	120.3	126.1	117.2	120.7	133.5
<i>SD</i>	41.3	35.4	40.4	40.8	38.1	47.4	33.5	42.8	43.7	51.5	43.1	35.1
Nonschizophrenics												
<i>M</i>	126.4	79.1	84.1	68.7	82.1	114.2	115.5	114.9	127.8	105.5	127.0	135.6
<i>SD</i>	34.3	48.3	48.5	50.0	54.8	59.1	29.3	47.0	41.2	55.4	44.6	41.9
Slow												
Schizophrenics												
<i>M</i>	112.5	25.7	20.7	19.6	21.0	15.9	108.1	40.9	38.1	34.0	39.0	34.0
<i>SD</i>	24.9	13.1	13.4	11.1	12.5	9.0	28.8	19.0	26.5	21.2	23.0	19.0
Nonschizophrenics												
<i>M</i>	112.2	17.7	16.1	13.7	15.1	12.6	115.1	17.4	14.8	12.5	13.2	11.3
<i>SD</i>	38.0	5.8	6.5	5.1	7.2	6.1	28.9	6.5	6.0	4.8	6.1	4.7

TABLE 3

REPEATED MEASURES ANALYSIS OF VARIANCE OF
TREATMENT EFFECTS ON TAPPING SCORES

Source	df	MS	F
Between subjects			
Subject groups (A)	1	4,057	0.73
Instructions (B)	1	696,737	126.15***
Delivery (C)	1	53,495	9.68**
A × B	1	5,571	1.01
A × C	1	3,923	0.73
B × C	1	24,583	4.45*
A × B × C	1	319	0.05
Error (B)	88	5,523	
Within subjects			
Trials (T)	5	49,694	132.94***
T × A	5	382	1.02
T × B	5	26,409	70.65***
T × C	5	3,103	8.30***
T × S × I	5	517	1.38
T × S × D	5	164	0.43
T × I × D	5	1,894	5.07***
T × S × I × D	5	667	1.78
Error (W)	440	374	

* $p < .03$.

** $p < .003$.

*** $p < .001$.

many trials to reach the criterion for "slower" than for "faster."

Experimental Data

Table 2 shows the mean scores for tapping for S and C groups for all conditions. A repeated-measures analysis of variance was performed on the data to evaluate the effects of counterbalancing the words (faster-slower and slower-faster). As there were no order effects for the C group ($F = 1.42$, $df = 1/40$) or for the S group ($F = 0.85$, $df = 1/40$) the data were combined for further analysis.

Three variables were considered in this analysis: two levels of pathology (schizophrenics and nonschizophrenics), two sources of delivery (Internal goad and External goad), and two pretask instructions (Minimal Instructions and Slow Instructions). The results of this analysis are shown in Table 3.

The table clearly shows that for both groups performance on the tapping task was modified as a result of the pretask instruction. For a Slow Instruction both the S and the C group significantly decreased tapping speed ($p < .001$). However, it was noted that there was also a modification in speed of

tapping which was determined by whether the goads were Internal or External ($p < .001$). The significant interaction of main effects, Instructions × Delivery, indicates qualifications of these effects, however.

To delineate more precisely the controlling factors involved in the interaction between Ss, and the number of significant interactions within Ss, Trials × Delivery, Trials × Instructions, and Trials × Instructions × Delivery, Duncan's multiple-range test was used to make relevant comparisons. The results will be described by noting the effects on the tapping performance of S groups of the treatments and then examining the effects of the goads on the groups. However, before describing these effects it must be noted that a comparison of the mean difference in operant levels of tapping performance for all groups and conditions was not significant.

Instruction and Delivery Effects on S and C Groups

Both S and C groups significantly inhibited speed of tapping for a pretask Slow Instruction ($p < .001$). However, while the schizophrenics tapped as slowly as the controls when self-verbalizing (Internal goad condition) this was not true for the External goad condition. Speed of tapping of schizophrenics under the External condition was significantly faster than the controls (probabilities ranging from .05 to .01 for all goad conditions). The difference was most pronounced when there was no External goad, that is, the "no word" condition ($p < .001$). The schizophrenics in the External condition did not tap significantly faster than the schizophrenics in the Internal goad condition, however. Further analysis showed this was due to the fact that the goad "slower" in each of these conditions did not reach an acceptable level of significance; all other goads were significantly separated. When the C groups were instructed to slow down, the source of delivery of goads, Internal or External, was not a factor influencing performance.

For the Minimal Instruction condition there was a significant decrement in speed of tapping from the operant level which occurred as a function of self-verbalization (Internal goad condition) for both groups ($p < .001$). No

such inhibition of tapping occurred in either of the S groups when the goads were External.

Word Meanings and Performance

The effects of the meanings of the goads on performance is best observed in the Minimal Instruction condition. Rate of tapping of schizophrenics for the goads "faster" and "slower" differ significantly for the Internal goad condition ($p < .05$), but not for the External goad condition. There was some inhibition of response for the goad "slower." However, this was only significant when compared with the upsurge of tapping that occurred for the "no word" condition ($p < .05$).

The goad difference for the C group in the Internal condition did not reach an acceptable level of significance. However, for the External condition these Ss responded to the meanings implied in the goads "faster" and "slower" by tapping faster and slower, respectively ($p < .02$), when the E spoke the words.

A pretask Slow Instruction was sufficient to eliminate the influence that the goad meanings themselves might have exerted over tapping for both the S and C groups ($p < .001$). Little or no fluctuation in tapping occurred as a function of the goads themselves when Ss had been given an instruction to tap slowly.

There is additional evidence that an instruction is superior to a goad in the inhibition of response. A comparison of the effectiveness of the goad "slower" for Minimal Instructions and Slow Instructions revealed that both S and C groups were better able to inhibit speed of tapping for the Slow Instructions ($p < .001$).

Rating Scale

A rating scale was used to determine whether the words "faster" and "slower" might have differing speed connotations for the two S groups. Pearson's chi-square of association was used to analyze the differences in rating these words by the two groups. There were no significant differences for the words "letter" or "slower." However, schizophrenics rated the word "faster" significantly more "rapid" than the C group ($\chi^2 = 26$, $p < .001$).

DISCUSSION

Relative to the questions asked concerning the parameters of verbal control of nonverbal behaviors of schizophrenics and nonschizophrenics the results of the present study indicated that: (a) Source of delivery of goads, Internal or External, is an important parameter of verbal control, but its effects are related to pretask instructions and to level of pathology. When no pretask instruction was given the externally delivered goads appeared to "instruct" the C group, but not the S group. It was noted, however, that this was true only for the External goad "slower," which implicitly demands response inhibition by the Ss. For the goad "faster," which demands no such inhibition, performance was practically the same for the two groups. The Internal goads had the interesting effect, for both groups, of markedly decreasing tapping speed from the operant level. There was also a significant decrease in tapping performance from the External goad condition. That is, while Ss were saying the goads they seemed unable to tap at the rapid pace which otherwise characterized their performance. When a pretask Slow Instruction was introduced both the S and C group were "instructed" in their performance. However, the degree to which schizophrenics were able to slow down their response, following the pretask instruction, was dependent upon source of goad delivery; the C group responded to the Slow Instruction in the same manner whether goads were Internal or External. Specifically, schizophrenics in the External goad condition tapped considerably faster than the schizophrenics in the Internal goad condition. The C group tapped at the same speed whether the goads were External or Internal. (b) The use of goads consistent, irrelevant, or contradictory to a pretask instruction is not a relevant dimension in verbal control of motor responses of schizophrenics or nonschizophrenics. While the External goads had a disinhibitory effect on speed of tapping of the schizophrenics, relative to other groups, there was no evidence that this disinhibitory effect was related to the verbal inconsistencies of the Instruction goad condition. There was little variation in speed of tapping for any of the goad condi-

tions whether these were consistent, irrelevant, or contradictory to pretask instructions.

The similarity of schizophrenics and non-schizophrenics for the Internal goad condition, in the present study, suggests a reevaluation of the role of external verbalizations and, possibly, external stimuli, in the modification of the behavior of schizophrenics. Paradoxically, while schizophrenics have been shown to improve their performance when given distinctive external cues concerning response performance (Lang, 1959), these patients are also characterized by their poor ability to utilize external cues in the modification of performance. Following a review of the literature, Buss and Lang (1965) report that on verbal conditioning tasks schizophrenics were consistently less able than normals to make use of verbal cues, for example, "right" or "wrong," to improve their performance. Letchworth (1966) reports that even with a pretask "focusing" instruction, which is known to facilitate the subsequent conditioning of normals, schizophrenics appear unable to make use of external verbal conditioning cues.

The findings that schizophrenics perform as well as normals for the Internal goad condition suggests that the most effective method by which to modify maladaptive behaviors of schizophrenics might be one which utilizes "self-speech" to produce voluntary control of responses. However, while "self-speech" may be beneficial in providing the schizophrenic with a means of inhibiting responses at will, it would scarcely have generality unless a concomitant ability to respond to the external cues was also acquired. The relevant differences which characterized the S and the C groups related to the ability of the C group to be guided by External goads in the inhibitory control of response. Conceivably, to develop in schizophrenics a self-instructional control which takes cues for control from external sources it might be necessary to condition the external control cue to the "self-speech" of the patient.

In this respect, it is of interest that performance of schizophrenics, in the present study, for the "no word" condition of the Slow Instruction treatment was markedly similar to tapping speeds in the goad condi-

tions preceding it. Schizophrenics inhibiting response under the Internal goad apparently acquired a self-instructional control available to them in the "no word" condition.

The failure to find that goads which were irrelevant or contradictory to the instruction disrupted performance of the schizophrenics on the tapping response is of interest. There is a theoretical formulation of schizophrenia, the double bind theory (Bateson, Jackson, Haley, & Weakland, 1956), which hypothesizes that the schizophrenic disorder is the end product of a faulty communication system which pulls the patient in opposing directions as he is required to respond to inconsistent and contradictory messages conveyed to him by a dominant figure or figures in his life. The present study provides no evidence that schizophrenics, or nonschizophrenics, are disturbed in their performance by the conflicting messages they receive. However, this study alone cannot be taken as an indictment of the double bind theory. The theory assumes that on one level the message to the schizophrenic involves emotional rejection. The design of the present study made it impossible to evaluate such components of the verbal message. However, as Schuham (1967), following a review of the literature, has concluded, there is little or no empirical support for the hypothesis, at present the very existence of the phenomenon appears open to serious question.

The ability of the schizophrenics, in the present study, to maintain a response set, unaffected by irrelevant stimulus elements, is contrary to results reported by other investigators (Payne & Caird, 1967; Rodnick & Shakow, 1940; Shakow, 1962). The hypothesis of a failure in response set put forward by these investigators has been based primarily upon RT tasks in which S is penalized for slight delays in responding to irregularly placed, or unusually distracting, stimuli. The differences in the results obtained in the present study and those of others may be due to the nature of the tasks which were used, or to the pretask instruction. The most plausible explanation would appear to be that the pretask instruction overcame the "set" deficit described by Shakow (1962).

Why the schizophrenics were less able to

inhibit response for the externally delivered goads than for the internally delivered goads is not immediately clear. It appears to this author that this deficit behavior represents a failure in self-control which occurred because these patients were less able to use the external discriminative stimulus to control responses. (When no element of *control* was involved, that is, no inhibition of response, as for the goad "faster," the schizophrenics responded as the controls.) Possibly, when schizophrenics were told to tap as slowly as they could and then heard *E* saying the goad "letter," the task was more confusing than when *Ss* verbalized the goad "letter." Such an interpretation would be in agreement with the position taken by Lang and Buss (1965) that schizophrenic deficit performance is an increasing function of task ambiguity or complexity.

However, it could be argued that the differences observed between the groups for the Internal and External conditions were due to a confounding of the frequency of goad presentation. Goads presented at a rapid rate could, conceivably, serve to confuse *S* in his assessment of task requirements. It will be recalled that the number of goads spoken by *Ss* in the Internal goad condition determined the frequency with which the external goad was presented. The mean frequency of the goad "faster" for schizophrenics was 57 spoken words for 30 seconds, for the control group it was 45.5; for the goad "slower" the mean was 46.5 for the schizophrenics and 36.5 for the control group.

There are several arguments which mitigate against the possibility that the differences between the two groups were due to frequency of goad presentations. First, the difference score between "faster" and "slower" was not different for the *S* group (mean difference = 11) and the *C* group (mean difference = 9.1). Second, the pilot data indicated a similar inability of schizophrenics to vary a tapping response according to the meaning of the External goads when the goads had been paced at a constant 2-second interval for schizophrenics and controls alike. Finally, under the Slow Instruction condition there was little difference in the frequencies of the goads for the schizophrenic and control

groups. (The means for "faster" were 23.9 for the schizophrenic group and 20.2 for the control group; for "slower" the mean was 17.5 for the schizophrenic and 15.8 for the control group. These differences were not significant.) If the *S* group had tapped at a rate which corresponded with the number of External goads presented there should have been no differences between this group and the *C* group. It was noted that the *C* group never tapped at a faster rate than the rate of External goad presentation.

While it could be argued that the present findings indicate, experimentally, the autistic style of control characteristic of schizophrenics, analogous findings with children have shown that self-verbalizations exert control over subsequent responses on concept reversal tasks, normally beyond the abilities of these children (Silverman, 1966). With the schizophrenics, as well as with children, two aspects of self-verbalization, a focusing set and increased information from response mediated feedback, may interact to provide the inhibitory control necessary for appropriate response. The present study indicates the need for research into ways of influencing the "self-speech" of schizophrenics.

REFERENCES

- BATESON, G., JACKSON, D. D., HALEY, J., & WEAKLAND, J. H. Toward a theory of schizophrenia. *Behavioral Science*, 1956, 1, 251-264.
- BIRCH, D. Verbal control of non-verbal behavior. *Journal of Experimental Child Psychology*, 1966, 4, 266-273.
- BUSS, A. H., & LANG, P. J. Psychological deficit in schizophrenia: I. Affect, reinforcement, and concept attainment. *Journal of Abnormal Psychology*, 1967, 70, 2-24.
- LANG, P. J. The effect of aversive stimuli on reaction time in schizophrenia. *Journal of Abnormal and Social Psychology*, 1959, 59, 263-268.
- LANG, P. J., & BUSS, A. H. Psychological deficit in schizophrenia: II. Interference and activation. *Journal of Abnormal Psychology*, 1965, 70, 77-106.
- LETCHWORTH, G. E. Studies in efficiency: Verbal conditioning in schizophrenia. *Journal of Consulting Psychology*, 1966, 30, 45-50.
- LOVAAS, O. Cue properties of words: The control of operant responding on rate and content of verbal operants. *Child Development*, 1964, 35, 245-256.
- LURIA, A. R. *Higher cortical functions in man*. New York: Basic Books, 1966.
- PAYNE, R. W., & CAIRD, W. K. Reaction time, distractibility, and overinclusive thinking in psy-

- chotics. *Journal of Abnormal Psychology*, 1967, 72, 112-121.
- RODNICK, E. H., & SHAKOW, D. Set in the schizophrenic as measured by a composite reaction time index. *American Journal of Psychiatry*, 1940, 97, 214-225.
- SCHUBAM, A. I. The double-blind hypothesis a decade later. *Psychological Bulletin*, 1967, 68, 409-416.
- SHAKOW, D. Segmental set: A theory of the formal psychological deficit in schizophrenia. *Archives of General Psychiatry*, 1962, 6, 17-33.
- SILVERMAN, I. Effect of verbalization on reversal shifts in children: Additional data. *Journal of Experimental Child Psychology*, 1966, 4, 1-8.
- ULLMANN, L. P., & GIOVANNI, J. The development of a self-report measure of the process-reactive continuum. *Journal of Nervous and Mental Disease*, 1964, 138, 38-42.
- WIENCKOWSKI, L. A. Stimulus factors influencing disjunctive reaction time in schizophrenic and "normal" subjects. Unpublished doctoral dissertation, University of Buffalo, 1959.
- ZIGLER, E., & PHILLIPS, L. Social competence and the process-reactive distinction in psychopathology. *Journal of Abnormal and Social Psychology*, 1962, 65, 215-222.

(Received May 8, 1968)

(Continued from page 225)

- Females' Affective Responses to Reading Erotic Literature: Donald L. Mosher* and Irene Greenberg: Department of Psychology, University of Connecticut, Storrs, Connecticut 06268.
- The Internal-External Control Dimension as Related to Accident and Suicide Proneness: Christene Blanton Williams* and James B. Nickels: Department of Psychology, University of Manitoba, Winnipeg, Canada.
- Adjustment, Mental Health Opinions, and Proficiency of Child Care Personnel: John W. Giebink* and Donald O. Stover: The Psycho-Educational Clinic, University of Wisconsin, Madison, Wisconsin 53706.
- Therapist Directiveness, Client Dominance, and Therapy Resistance: Robert W. Hagebak and George V. C. Parker*: Department of Psychology, University of Texas, Austin, Texas 78712.
- Temporal Correlates of the Achievement Value and Manifest Anxiety: Thomas J. Cottle*: Department of Social Relations, Harvard University, Cambridge, Massachusetts 02138.
- Sensation Seeking, Impulsivity, and Psychopathic Personality: R. Blackburn*: Department of Psychology, Broadmoor Hospital, Crowthorne, Berkshire, England.
- Locus of Control in Psychiatric Patients: Martin Harrow* and Anthony Ferrante:
- School of Medicine, Yale University, 333 Cedar Street, New Haven, Connecticut.
- Study of Psychological Content in the Communications of Suicidal Individuals: Allen R. Darbonne*: Veterans Administration Center, Wilshire and Sawtelle Boulevards, Los Angeles, California 90073.
- Comparison of the Binet and WPPSI with Lower-Class Five-Year-Olds: Joen Fagan* Elizabeth Broughton, Mildred Allen, Betty Clark, and Patricia Emerson: Department of Psychology, Georgia State College, 33 Gilmer St., S. E., Atlanta, Georgia 30303.
- Participation in Psychological Research: Relation to Birth Order and Demographic Factors: Edward H. Fischer* and David Winer: Department of Psychology, Connecticut Valley Hospital, Middletown, Connecticut 06457.
- Direction of Affective Influence in Psychotherapy: Kenneth I. Howard,* Merton S. Krause, and David E. Orlinsky: Institute for Juvenile Research, 232 East Ohio Street, Chicago, Illinois 60611.
- Prediction Enhancement through the Use of Moderator Variables: James E. Cowden*: Department of Health and Social Services, Division of Corrections, 1 West Wilson Street, Madison, Wisconsin 53701.
- Intelligence Defined as Effectiveness of Approaches: Constance T. Fischer*: Psychology Department, Duquesne University, Pittsburgh, Pennsylvania 15219.

* Asterisk indicates author for whom address is supplied.

(Continued on page 239)

SELF-CONCEPT OF THE MARRIED PSYCHIATRIC PATIENT AND HIS MATE'S PERCEPTION OF HIM

MARTIN HARROW,¹ DAVID A. FOX, AND THOMAS DETRE

Yale University

Twenty-six married psychiatric patients and their spouses were administered a 50-item questionnaire to assess their concepts of themselves and of each other 1 week after acute hospitalization and then again 7 weeks later. The patients' initial self-images and their mates' perception of them were both more negative than their own and their spouses' images of the average person ($p < .001$). The patients and their mates viewed the patient as the most disturbed family member, with both seeing the spouse in a more positive light than the patient ($p < .01$). The patients' self-images improved during hospitalization ($p < .01$). The spouses' images of the patients also tended to become more positive ($p < .10$), but the spouses still viewed the patients more negatively than themselves.

Despite the increasing emphasis on the self-concept in the recent literature (Wylie, 1961), there are relatively few systematic studies of the self-perception of hospitalized adult psychiatric patients and of changes in their self-perception during hospitalization. In addition, despite the current popularity of family treatment and of theories about the role of the family in psychopathology (Ackerman, 1966), there has been even less research on how family members perceive the hospitalized psychiatric patient.

Our previous research with hospitalized adolescent patients and their families indicated that despite frequent family discord both patients and their parents viewed each other in a way that was generally in accord with the views of society (i.e., the patients were seen in a negative light and the other family members were viewed in a more positive manner) (Harrow, Fox, Markhus, Stillman, & Hallowell, in press). The current research attempts to determine whether this trend is also true of marital families.

The specific questions which the present study attempts to answer are:

1. What is the self-perception of the married psychiatric patient shortly after his admission to a psychiatric hospital?

2. Do patients and/or their mates see only the patient in a negative light (or as the most

disturbed family member), or do they see both the spouse and the patient negatively?

3. Does the patient's self-concept change during the course of hospital treatment, and do other family members' (e.g., his spouse's) perceptions of the patient change?

METHOD

Subjects and Setting

The Ss were 26 married psychiatric inpatients and their husbands or wives. The mean age of the patient sample was 37.5 years and the mean educational level was 14 years. There were 9 males and 17 females among the patient population. Diagnostically the patients included 10 schizophrenics (38%), 10 depressives (38%), 4 character disorders or neurotics (15%), and 2 of varied diagnoses (8%).

The patients were all recent admissions to the Acute Psychiatric Inpatient Division of Yale-New Haven Hospital. This ward, a short-term intensive treatment unit, emphasizes group, family, and social treatment (Astrachan, Harrow, & Flynn, 1968; Detre, Kessler, & Jarecki, 1963).

Procedure and Data Collection

The patients and their spouses were tested, individually, at two time periods: (a) Time Period 1 (TP1), a week and a half after hospitalization, and (b) Time Period 2 (TP2), 7 weeks later. Complete sets of data were collected at both TPs from 23 patients and their spouses, and partially complete data were collected from the other three patients and their mates due to the difficulty of testing both patients and their spouses at two TPs.

During pilot work the original deck of 100 Q-sort items used by Butler and Haigh (1954) was reduced to 50 to eliminate items that would be difficult to understand by patients with below-average intelligence. Pilot work also suggested advantages in

¹ Requests for reprints or a copy of the 50 items used in the test may be obtained from Martin Harrow, Department of Psychiatry, Yale University, 333 Cedar Street, New Haven, Connecticut 06510.

TABLE 1
MEAN ADJUSTMENT SCORES AT TIME PERIOD 1 AND TIME PERIOD 2

People rating	Concept rated	Mean adjustment scores		$t(M_{diff})$ — TP1 vs. TP2
		TP1	TP2	
Patients	Self	7.65	20.52	3.68**
Patients	Self \neq Ideal	$r = 0.24^a$	$r = 0.54^a$	4.24***
Patients	Ideal	31.00	37.61	2.20*
Patients	Average person	24.00	28.00	1.90
Patients	Spouse	20.22	26.78	2.77*
Spouses	Patient	12.78	18.17	1.73
Spouses	Ideal	37.30	38.96	1.83
Spouses	Self	23.04	26.13	1.35
Spouses	Average person	27.13	28.52	0.67
Normal controls	Self	26.20	—	—

Note.—TP1 = Time Period 1; TP2 = Time Period 2.

^a Represents mean correlation (using Fisher's transformation) between patients' self-images and their ideals for themselves.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

modifying the typical forced-choice distribution, due to difficulties encountered by disturbed patients and anxious mates in attempting to comply accurately with forced-choice instructions. The final test form was a 50-item paper-and-pencil test, with each item rated on 9-point Likert-type scales.

At each TP the patient and his mate rated each of the following concepts, separately, on the 50 items: (a) the patient, (b) the patient's spouse, (c) the ideal for the patient, and (d) the ordinary average person.

Each of the 50 items had been rated for "adjustment" and for social desirability in a previously reported study (Wiener, Blumberg, Segmen, & Cooper, 1959). Using these ratings, the 50 items (22 positive, 22 negative, and 6 neutral items) were combined and a total adjustment score was calculated for each image of each patient and his mate. Scores for each image could range from +44 to -44. The major part of the data consisted of 16 adjustment scores for each family (one for each image at each of the two TPs) and correlations of all images within each family. Ratings of family interaction and of the daily behavior and symptoms of all patients were collected from the patients' therapists and from other staff observers at each testing period.

RESULTS

Table 1 presents the mean "adjustment" scores at both TPs and t tests (M_{diff}) comparing the scores at TP1 with those at TP2. Table 2 reports the t tests comparing the key concepts with each other at TP1, and also comparing these key concepts with each other at TP2.

The initial adjustment scores derived from the self-images of the patients were significantly lower than their own and their spouses' images of the ordinary average person (Tables 1 and 2, $p < .001$). Their initial self-images were also significantly more negative than the self-images of a control group of 20 "normals" used in a previous study, employing the same test instruments (Tables 1 and 2, $p < .001$).

In addition, the patients' initial views of themselves were significantly more negative than their views of their spouses (Tables 1 and 2, $p < .01$). The spouses tended to agree with this, seeing themselves in a more favorable light than they saw their patient-mates (Tables 1 and 2, $p < .001$). Thus, both the patients and their spouses tended to see the spouse in a relatively positive light. As further evidence for this, the patients' spouses did not see themselves as significantly more negative than the average person (Table 2, $p > .10$), and the patients' views were in general agreement with this (Table 2, $p > .10$).

After 8 weeks of hospitalization the mean correlation between the patients' self-images and their ideals had become significantly higher (Table 1, $p < .001$). Furthermore, their self-images had become significantly more positive (Table 1, $p < .01$), but were still lower than both their own and their

TABLE 2
TESTS OF SIGNIFICANCE COMPARING THE KEY CONCEPTS^a AT TIME PERIOD 1
AND TIME PERIOD 2

People rating	Concept rated	vs. ^a	People rating	Concept rated	Time Period	
					1	2
Patients	Self	vs.	Normal Controls	Self	4.37***	1.31
Patients	Self	vs.	Patients	Average person	5.20***	2.23*
Patients	Self	vs.	Spouses	Average person	4.92***	2.29*
Patients	Self	vs.	Patients	Spouse	3.39**	1.58
Spouses	Patient	vs.	Spouses	Self	4.31***	2.50*
Patients	Spouse	vs.	Patients	Average person	1.09	0.33
Spouses	Self	vs.	Spouses	Average person	1.58	1.15

^a In all cases the mean adjustment score for the second concept listed is higher than that for the first concept listed.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

spouses' image of the average person (Table 2, $p < .05$).

The spouses' images of the patients also tended to become more positive 7 weeks later (Table 1, $p < .10$), but less than did the patients' own views of themselves. Furthermore, at TP2 the spouses *still* viewed the patient more negatively than themselves, and the patients showed some tendency to concur in this view (Table 2).

DISCUSSION

Initial Self-Concepts of Patients

The initially negative self-concepts of the patients are in marked contrast to the images of the average person held by both the patients and their spouses, and are significantly more negative than the self-concepts of the "normal" control group. In addition, when we compare the patients' initial views of themselves with their views of their mates, we also find a significant difference. Thus, these married *patients* see themselves as the more disturbed family member and this is in agreement with our results on hospitalized adolescent patients (Harrow et al., 1968).

The dramatic impact of acute hospitalization, with its social stigma (Lamy, 1966), may have influenced the patient's self-perception. Although hospitalization frequently places the patient in a more supportive setting than before, with some subsequent symptom

reduction (Colbert & Harrow, 1967; Harrow et al., 1966), the complete removal from the home and family also implies a collapse of the patient's previous adjustment. This implication of collapse probably lowers the patient's self-concept.

In addition, the duration of previous outpatient therapy by the patients was related significantly to their initial self-concepts ($r = -.45$, $N = 26$, $p < .05$), tentatively suggesting that more chronic patients who have been in treatment longer have more negative self-concepts.

Spouses' Initial Views of Patients and of Themselves

The patients' mates agreed with the patients' negative self-evaluations, although their initial judgments were not as harsh as were the patients'. In contrast, both marital partners consistently saw the nonhospitalized spouses in a more positive light. These results occurred despite sharp marital discord in some families, intense ambivalence and overinvolvement, and acting out by some patients and their mates towards each other. This negative view of the patient by another family member is also in agreement with data we have reported on the views of families of hospitalized adolescents.

The present results and those of our previous work contain some interesting implica-

tions, since almost all patients and families studied were seen in family treatment. While the patient and his family may pay lip service to the notion that in family treatment both parties are "patients," both of their moderately positive evaluations of the nonhospitalized mate suggest only a partial acceptance of the idea that both are patients.

Changes in Patients' Self-Concepts and Spouses' Views over Time

The patients' more positive self-concepts 7 weeks later seem to reflect their clinical improvement. In addition, the correlation of the patient's ideal with his view of himself also became significantly higher over time.

Since the patients' ideals tended to become more positive at *TP2*, the higher correlations of patients' self-images with their ideal images were not due to a lowering of ideal images (see Table 1), but may have been related to the patients' self-images becoming more positive at *TP2*. It should be noted in contradiction to some theories which postulate more realistic (and lower) patient ideals as clinical improvement occurs, the results with the present sample and two previous samples of ours suggest that as patients improve clinically and their self-images become higher, there is a slight trend for their ideals to become higher. Although this result may run counter to some theories, it would fit in with a model in which peoples' expectations become higher as their competence increases.

The spouses agreed with the more positive self-evaluations of the patients. At *TP2*, however, the spouses still viewed the patients more negatively than themselves ($p < .05$), and there was also some tendency for the patients to see themselves more negatively than they saw their spouses. This occurred despite the more positive image of the patients held by both parties. These results concur with the view of society at large, since the patient resides in a mental hospital and is still considered "sick," despite his considerable clinical improvement during the 7-week interim. The tendency for the spouses to view the patients more negatively than they see themselves, and the less optimistic view by the spouses concerning patients' improvement, may create subsequent difficulties for the pa-

tients after their discharge from the hospital, as there is evidence that relatives' attitudes and expectations bear some relation to patients' posthospital success (Freeman & Simmons, 1963).

In relation to changes in the patients' and spouses' views, the present authors computed, at each *TP*, correlations of patients' with spouses' views of the patients, and also correlations of patients' with spouses' views of the spouses. These correlations indicated that there was significantly greater agreement as to what the patients (and the spouses) are like at *TP2* as compared to *TP1* ($p < .02$). The greater agreement about what the patient is like may be related, in part, to a reduction in turmoil and conflict around the patient as time passes since acute hospitalization, and as the patient improves. The data indicating greater agreement about the spouse may also suggest some reduction in conflict concerning the spouse (who either assisted or was involved in the hospitalization of the patient) as time passes. In addition, the present hospital setting emphasizes family treatment, with patients and families being involved in conjoint family therapy (Satir, 1964) and multiple family group therapy (Harrow, Astrachan, Becker, Detre, & Schwartz, 1967), and this may have played some role in the greater family agreement about what the patient and spouse are like at *TP2*.

REFERENCES

- ACKERMAN, N. W. Family therapy. In S. Arletti (Ed.), *American handbook of psychiatry*. Vol 3. New York: Basic Books, 1966.
- ASTRACHAN, B. M., HARROW, M., & FLYNN, H. R. Influence of the value-system of a psychiatric setting on behavior in group therapy meetings. *Social Psychiatry*, 1968, 3, 165-172.
- BUTLER, J., & HAIGH, G. Changes in the relation between self-concepts and ideal concepts consequent upon client-centered counseling. In C. Rogers & R. Dymond (Eds.), *Psychotherapy and personality change*. Chicago: University of Chicago Press, 1954.
- COLBERT, J., & HARROW, M. Psychomotor retardation in depressive syndromes. *Journal of Nervous and Mental Disease*, 1967, 145, 405-419.
- DETRE, T. P., KESSLER, D. R., & JARECKI, H. G. The role of the general hospital in modern community psychiatry. *American Journal of Orthopsychiatry*, 1963, 33, 690-700.

- FREEMAN, H. E., & SIMMONS, O. G. *The mental patient comes home*. New York & London: Wiley, 1963.
- HARROW, M., COLBERT, J., DETRE, T., & BAKEMAN, R. Symptomatology and subjective experiences in current depressive states. *Archives of General Psychiatry*, 1966, 14, 203-212.
- HARROW, M., ASTRACHAN, B. M., BECKER, R. E., DETRE, T., & SCHWARTZ, A. H. An investigation into the nature of the patient-family therapy group. *American Journal of Orthopsychiatry*, 1967, 37, 888-899.
- HARROW, M., FOX, D. A., MARKHUS, K. L., STILLMAN, R., & HALLOWELL, C. B. Changes in adolescents' self-concepts and their parents' perceptions during psychiatric hospitalization. *Journal of Nervous and Mental Disease*, 1968, 147, 252-259.
- LAMY, R. E. Social consequences of mental illness. *Journal of Consulting Psychology*, 1966, 30, 450-455.
- SATIR, V. *Conjoint family therapy*. Palo Alto: Science and Behavior Books, 1964.
- WIENER, M., BLUMBERG, A., SEGMENT, S., & COOPER, A. Judgment of adjustment by psychologists, psychiatric social workers, and college students, and its relationship to social desirability. *Journal of Abnormal and Social Psychology*, 1959, 59, 315-321.
- WYLIE, R. C. *The self-concept*. Lincoln: University of Nebraska Press, 1961.

(Received May 13, 1968)

(Continued from page 234)

- Operant-Interpersonal Treatment for Marital Discord: Richard B. Stuart*: University of Michigan, School of Social Work, 1065 Frieze Building, Ann Arbor, Michigan 48104.
- Verbal and Overt-Behavioral Assessment of a Specific Fear: Anthony F. Fazio*: Department of Psychology, University of Wisconsin, Milwaukee, Wisconsin 53201.
- Rorschach Prediction of Success in Clinical Training: A Second Look: Rae Carlson*: Division of Psychological Studies, Educational Testing Service, Princeton, New Jersey.
- Effect of Information, Subject Age, and Instructional Set on Predictive Judgments: James L. Knestrick* and Leon Gorlow: 270-E North Arlington Avenue, Harrisburg, Pennsylvania 17109.
- Personal Adjustment as a Determinant of Aggressive Behavior toward the Mentally Ill: Kenneth Ring* and Amerigo Farina; Department of Psychology, University of Connecticut, Storrs, Connecticut 06268.
- Relations between Personality and Smoking Behavior in Preadult Subjects: Gene M. Smith*: Department of Anaesthesia, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts 02114.
- Effects of Cues and Examiner Influence on Two Wechsler Subtests: Jerome M. Sattler*: Division of Life Sciences, San Diego State College, San Diego, California 92115.
- Personality Correlates of Accurate Empathy in a College Companion Program: George J. Vesprani*: Hamilton County Diagnostic Clinic for the Mentally Retarded, 295 Erkenbrecher Avenue, Cincinnati, Ohio 45229.
- Some Personality and Aptitudinal Correlates of the "A-B" Therapist Scale*: James E. Dublin,* Charles F. Elton, and Juris I. Berzins: Counseling and Testing Center, University of Kentucky, Lexington, Kentucky 40506.
- SPI Discriminators among Four Psychological Health-Sickness Levels: Donald E. Spiegel*: Veterans Administration Center, Wilshire and Sawtelle Boulevards, Los Angeles, California 90073.
- Abstraction on Proverbs in Process-Reactive Schizophrenia: Robert Cancro*: The Menninger Foundation, Topeka, Kansas 66601.

* Asterisk indicates author for whom address is supplied.

(Continued on page 253)

EFFECTS OF ANXIETY ON TIMED AND UNTIMED INTELLIGENCE TESTS:

ANOTHER LOOK¹

LARRY W. MORRIS² AND ROBERT M. LIEBERT

Vanderbilt University

The timed subtests of the WAIS were administered to 48 Ss, only half of whom were aware of being timed. The Ss were divided into four groups on the basis of their response to 15 worry items and 15 emotionality items taken from the Taylor Manifest Anxiety scale. Low-worry Ss who were "timed" were superior to those who were "untimed," whereas high-worry Ss tended to perform more poorly under the "timed" than the "untimed" condition. This Worry \times Timing interaction closely parallels the "Anxiety" \times Timing interactions found by others. Emotionality was not found to interact with the task variables.

The major formulations of the relationship of anxiety to performance (Mandler & Sarason, 1952; Spence, 1958) have led to an examination of various aspects of performance situations, some of which have differential effects on high- and low-anxious Ss. One factor to which attention has been directed in this regard is the *timing* of performance. Thus, Matarazzo, Ulett, Guze, and Saslow (1954) found anxiety, measured by Taylor Manifest Anxiety (MA) scale scores, to be negatively related to scores on a timed intelligence test, but unrelated to scores on untimed measures of intelligence. Similarly, Siegman (1956) administered the MA scale and the Wechsler Adult Intelligence Scale (WAIS), compared the performance of the high- and low-anxious Ss on the timed and untimed subtests, and found that "only the high MAS scorers had significantly lower scores on the time-limited subtests of the WAIS [p. 177]." And, in an experimental study, Sarason, Mandler, and Craighill (1952) found high- and low-anxious Ss, identified by Test Anxiety Questionnaire (TAQ) scores, to be differentially affected by different instructions involving the timing of a task. Specifically, Sarason et al. gave college

students a task which was impossible to finish within the allotted time. Half of these Ss were told that they were expected to finish; the rest were told that they were not expected to finish. Low-anxiety Ss were found to perform better under the stressful (i.e., "expected to finish") instructions than under the non-stressful condition. Differences in the performance of high-anxious Ss were slight, but, as predicted, they tended to perform less well under stress than otherwise.

In an attempt to refine existing conceptualizations, the present authors have recently questioned the value of viewing anxiety as a unitary state and have proposed an empirical examination of the component feelings and cognitions which are popularly called "anxiety" (Liebert & Morris, 1967). Specifically, we posited a distinction between cognitive and physiological-affective anxiety components. Thus, worry (*W*) was identified as cognitive concern about one's performance, while emotionality (*E*) referred to autonomic reactions which occur in response to the stress of the test-taking situation. Within this framework it has been demonstrated that (a) worry varies as an inverse function of one's performance expectancy in a test situation, while emotionality is unrelated to expectancy; and (b) emotionality rises sharply immediately before the test and decreases sharply immediately after leaving the test situation, whereas worry reaches a high level earlier and does not dissipate so rapidly (Liebert & Morris, 1967; Spiegler, Morris, & Liebert, 1968).

¹ This research was supported by a United States Public Health Service Predoctoral Fellowship, No. MH-34053, from the National Institute of Mental Health. The authors wish to express appreciation to Richard King for serving as experimenter in this study.

² Requests for reprints should be sent to Larry W. Morris, Department of Psychology, Vanderbilt University, Nashville, Tennessee 37203.

The present study attempted to apply the worry-emotionality distinction to an examination of the relationship of anxiety to intelligence test performance. We anticipated a partial replication of the findings cited above with respect to a Timing \times Anxiety interaction and advanced two additional hypotheses derived from the worry-emotionality distinction. Manipulation of task difficulty was also included (*a*) because this dimension has often been found to interact with other task manipulations in their effect on the performance of high- and low-anxious Ss (e.g., Sarason & Palola, 1960), and (*b*) because difficulty is closely related to performance expectancy, which has been used in prior worry-emotionality studies.

The major hypotheses of this study were as follows:

1. High-anxious Ss should do better on untimed than on timed tests, whereas low-anxious Ss should do better on timed than on untimed tests. This is based on the findings cited above.

2. (*a*) Timing of the test was conceived of as an external, stressful aspect of the testing situation, and was thus expected to elicit primarily emotionality. (*b*) Assuming emotionality to be inhibitory in nature, we would therefore expect high-emotionality Ss to perform poorly on timed tests relative to untimed tests. Like the low-anxiety Ss, low-emotionality Ss should perform better on timed than on untimed tests.

3. (*a*) Presentation of a difficult test was expected to lower performance expectancy, thus causing worry to increase. (*b*) Assuming worry also to be inhibitory in nature, we therefore expected that high-worry Ss would be inferior to low-worry Ss on hard tests, but superior to them on easy ones.

With respect to Hypotheses 2 and 3, it was assumed that the predicted differential effects of the anxiety components would occur only if the manipulations of timing and difficulty were effective in eliciting emotionality and worry, respectively. In other words, if Parts *a* of these hypotheses were not confirmed, then this experiment would be unable to provide a test of Parts *b*.

METHOD

Subjects

The Ss were 48 students in introductory psychology at Vanderbilt University, all of whom volunteered in order to receive the experimental credit required for completion of the course. They were grouped into four anxiety groups (high worry-high emotionality—HH; high worry-low emotionality—HL; low worry-high emotionality—LH; low worry-low emotionality—LL) on the basis of their response to an anxiety questionnaire administered to all students (approximately 350) early in the semester. This "Experiences Questionnaire" consisted of 15 worry items and 15 emotionality items, all of which were taken from the *MA* scale. On the basis of independent judging by the authors, items were chosen for which there was agreement that they reflected either cognitive concern or autonomic arousal. Students responded on the basis of a 5-point scale: "The statement is never true, seldom true, sometimes true, mostly true, or true for me." *W* and *E* scales were considered separately, high and low scores being defined by the upper and lower thirds on each scale. Students falling into the anxiety categories outlined above were then allowed to volunteer as Ss.

Materials

The performance task which was presented to Ss consisted of five subtests of the WAIS which are normally timed subtests (arithmetic, picture completion, block design, picture arrangement, and object assembly). Unlike Siegman (1956), who compared Ss' performance on the timed versus the untimed subtests of the WAIS, the present study compared the performance of Ss who were aware of being timed versus Ss who were unaware of being timed on identical subtests. This latter procedure appears to reduce the likelihood of contamination by task differences. The difficulty manipulation was accomplished by choosing items from the first part of the subtests (easy) versus items from the last part of each subtest (hard).

Procedure

The Ss entering the experimental situation were told that the aim of the experiment was to investigate the relationships between grades, feelings and attitudes, and IQ. They were told that they were selected as potential Ss on the basis of their present grade in introductory psychology. They were further told that they were going to be tested on a part of the Wechsler Adult Intelligence Scale, and that their scores could be transformed into IQ scores, thus providing a measure of their intelligence. Those in the *Timed* condition were told that the tasks would be timed, and that the time taken to reach a solution would be taken into account in determining their score. In the *Untimed* condition, nothing was said about time limits.

The Ss were then asked to complete a pre-examination questionnaire consisting of five *W* items

TABLE 1
ANALYSES OF VARIANCE

Source	WAIS scores		Emotionality change scores		Worry change scores	
	<i>MS</i>	<i>F</i>	<i>MS</i>	<i>F</i>	<i>MS</i>	<i>F</i>
Difficulty (A)	1140.75	18.29****	9.19	2.31	46.02	6.42**
Emotionality (B)	44.08	<1	1.69	<1	.52	<1
Worry (C)	1.33	<1	2.52	<1	2.52	<1
Timing (D)	252.08	4.04*	.02	<1	1.02	<1
A × B	.09	<1	1.69	<1	.02	<1
A × C	48.00	<1	.19	<1	.02	<1
A × D	420.09	6.74**	2.52	<1	.52	<1
B × C	65.34	1.05	4.69	1.18	.52	<1
B × D	24.09	<1	2.52	<1	22.69	3.17
C × D	432.01	6.93**	11.02	2.77	7.52	1.05
A × B × C	16.33	<1	7.52	1.89	2.52	<1
A × B × D	2.07	<1	2.52	<1	13.02	1.82
A × C × D	833.32	13.36****	.02	<1	7.52	1.05
B × C × D	47.99	<1	1.02	<1	1.02	<1
A × B × C × D	61.21	<1	.52	<1	9.19	1.28
Subjects/groups	62.38		3.98		7.17	

* $p < .06$.
 ** $p < .02$.
 *** $p = .001$.
 **** $p = .0002$.

and five *E* items taken from Mandler and Sarason's TAQ (cf. Liebert & Morris, 1967). The five subtests of the WAIS were then administered according to standard procedures, with the exception of differential timing procedures. In the Timed condition, the experimenter held a stop watch in his hand, making an obvious effort to time *S*'s work accurately. In the Untimed condition, the tasks were timed *without S's knowledge* by the use of a silent clock with a sweep hand, which was hidden from *S*'s view. Following completion of the task, *Ss* completed a postexamination questionnaire consisting of the same 10 items from the TAQ. Finally, *Ss* were assured that IQ scores could not be obtained from these tests alone, and that the personal information was strictly confidential.

Two dependent measures were used: (a) scores on the WAIS subtests, obtained by standard scoring procedures, and (b) *W* and *E* change scores, obtained by subtracting *W* and *E* scores on the preexamination questionnaire from the corresponding scores on the postexamination questionnaire.

RESULTS AND DISCUSSION

Analysis of variance of WAIS scores is presented in Table 1. It is evident here that, as expected, there is a main effect for difficulty level. There is also an effect for timing that approaches significance, performance on timed tests being better than on untimed tests. Additionally, there is an interaction effect of Difficulty × Timing, due to inferior perform-

ance by the *Ss* who were given hard, untimed tests. It is also clear from Table 1 that *W* interacts with the task variables of Timing and Difficulty × Timing, while *E* is involved in no significant interactions.

Table 1 also presents analyses of variance of *W* and *E* change scores. The only significant effect is that of difficulty level on *W* change scores, which primarily reflects a decrease in *W* scores when an easy test was presented. There was also a tendency, though not significant, for *W* scores to increase as predicted when a hard test was presented ($t = .73$). Contrary to prediction, *E* scores did not increase as a function of timing the test. In fact, there were no significant effects in the analysis of variance of *E* change scores.

The major hypotheses of this study can be evaluated in the light of the above data as follows:

Hypothesis 1. High and low anxiety was defined by presence in the HH or LL group, respectively. Figure 1a shows the Anxiety × Timing interaction obtained by Siegman (1956), and Figure 1b shows the corresponding findings in this study. Here it is evident that the same pattern of results was obtained in both investigations. However, the difference

between performance on timed versus untimed tests is not statistically significant for either high- or low-anxious Ss in this study, though the tendency in both cases is in the predicted direction ($t = 1.11$ and $t = 1.04$, respectively). Likewise, an analysis of variance of WAIS scores revealed a nonsignificant Anxiety \times Timing effect ($F = 2.03$, $p = .17$). There is, however, a significant three-way interaction ($F = 7.39$, $p < .02$), which indicates that the Anxiety \times Timing interaction is more pronounced for hard tests than for easy ones.

Hypothesis 2. This hypothesis received only partial support from the data: (a) the timing procedure failed to elicit emotionality, and (b) since this was prerequisite to the prediction concerning the performance of high-*E* Ss, the latter prediction could not be evaluated. Actually, WAIS scores for both high- and low-*E* Ss were ordered according to the prediction for the latter, their performance being better on timed than on untimed tests (Figure 1d). This finding seems reasonable, given the failure of the timing procedure to elicit emotionality.

Hypothesis 3. (a) The first part of this hypothesis received partial support, since the *W* scores of Ss who were given hard tests increased significantly more than the *W* scores of those who were given easy tests. However, this difference was largely due to the decreased *W* scores of the latter, rather than an increase upon presentation of the hard test. It is possible that the initial instructions led Ss to expect a hard test, so that those who were presented with an easy test experienced considerable relief from worry. On the other hand, presentation of a hard test aroused only a small amount of additional worry. Nevertheless, this finding further corroborates the inverse relationship found between expectancy and worry scores in earlier studies. (b) There is little evidence to indicate that the different levels of worry had differential effects on performance. Worry does not interact with difficulty alone; it does, however, enter into a three-way interaction with difficulty and timing, the two-way interaction between the task variables being accentuated for low-worry Ss.

The most significant finding in this study appears to be the interaction of *W* \times Timing.

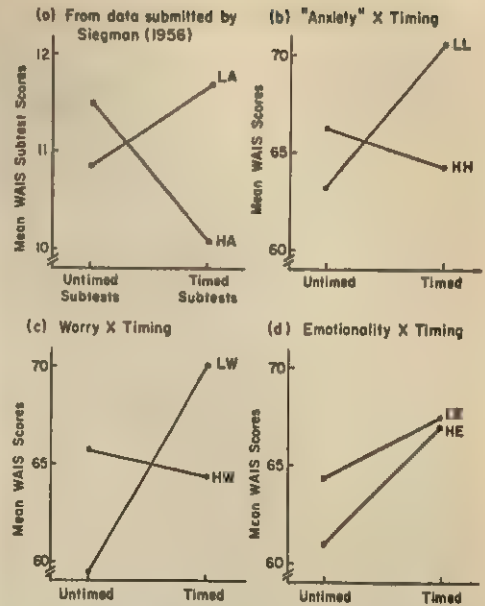


FIG. 1. Mean WAIS scores as a function of Timing and MA scale, Worry, and Emotionality.

Figure 1c portrays this interaction graphically. It is important to note that this interaction follows the same pattern as the "Anxiety" \times Timing interaction obtained in this and previous studies. Thus, these findings concerning worry are almost identical to those of Sarason et al. (1952) using "anxiety." In both cases, there was a clear difference in the performance of low-anxiety (worry) Ss under different conditions, but only a tendency for the performance of high-anxiety Ss to differ as predicted. These findings appear to lend very strong support to the distinction between worry and emotionality, and to the usefulness of this distinction in anxiety-performance research.

The evidence presented here in conjunction with the findings of previous studies clearly confirms the hypothesis that worry varies in a test situation as a function of the examinee's expectation concerning his performance on the test. Thus, when a difficult task is presented, S's expectation decreases, causing him to be more worried (i.e., concerned about his performance). The findings of the present study further suggest that it is worry, not "anxiety," which affects performance on intellectual-cognitive tasks and which inter-

acts with the relevant variables of the test situation.

On the other hand, emotionality scores have been found to vary as a function of being in or out of a test situation. However, the failure of the Timed condition to elicit an increase in emotionality stresses a problem which requires further research, that is, the problem of determining what specific aspects of a testing situation are principally responsible for the elicitation of emotionality. It also seems likely that there are some types of tasks (e.g., motor skills) in which emotionality may be the component which affects performance. Further experiments are planned to attempt to address these issues.

In contrast to the data of Siegman (1956), which suggested a "simple" interpretation concerning the inhibitory or disruptive effect of high anxiety on timed tests, the findings of this study are very similar to those of Sarason et al. (1952) in that they require consideration of differential effects of a stressor (timing) on high- and low-anxiety (or worry) Ss. Thus, timing the test seems to have an inhibitory effect on high-anxious Ss, and a facilitating effect on low-anxious Ss. In this study, the differential performance of low-worry Ss on timed and untimed tests seems best explained by assuming that the S with a low worry score is one who requires additional external motivation in order to perform well on this kind of task. The Timed condition is such a motivating condition, which enables the low-worry S to realize his potential. High-worry Ss are also motivated additionally by the Timed condition but, according to Mandler and Sarason (1952), have learned task-irrelevant anxiety responses which distract them from the task and thus disrupt their performance. Further studies are

needed to test this and other possible interpretations, and to further clarify the nature of the effect of both worry and emotionality on performance in various types of situations.

Overall, the aim of this experiment has been realized: that is, to show that the worry-emotionality distinction is applicable to analyses of the relationship of anxiety and performance and to suggest that by attending to these components of anxiety, we may be better able to specify the effects of anxiety on different kinds of performance tasks.

REFERENCES

- LIEBERT, R. M., & MORRIS, L. W. Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, 1967, 20, 975-978.
- MANDLER, G., & SARASON, S. B. A study of anxiety and learning. *Journal of Abnormal and Social Psychology*, 1952, 47, 166-173.
- MATARAZZO, J. D., ULETT, G. A., GUZE, S. D., & SASLOW, G. The relationship between anxiety level and several measures of intelligence. *Journal of Consulting Psychology*, 1954, 18, 201-205.
- SARASON, I. G., & PALOLA, E. G. The relationship of test and general anxiety, difficulty of task, and experimental instructions to performance. *Journal of Experimental Psychology*, 1960, 59, 185-191.
- SARASON, S. B., MANDLER, G., & CRAIGHTILL, P. G. The effect of differential instructions on anxiety and learning. *Journal of Abnormal and Social Psychology*, 1952, 47, 561-565.
- SIEGMAN, A. W. The effect of manifest anxiety on a concept formation task, a nondirected learning task, and on timed and untimed intelligence tests. *Journal of Consulting Psychology*, 1956, 20, 176-178.
- SPENCE, K. W. A theory of emotionally based drive (D) and its relation to performance in simple learning situations. *American Psychologist*, 1958, 13, 131-141.
- SPIEGLER, M. D., MORRIS, L. W., & LIEBERT, R. M. Cognitive and emotional components of test anxiety: Temporal factors. *Psychological Reports*, 1968, 22, 451-456.

(Received May 13, 1968)

PREFERENCE FOR COMPLEXITY IN ACUTE SCHIZOPHRENICS

JOHN M. NEALE¹ AND RUE L. CROMWELL

Vanderbilt University

A paired comparisons scaling technique was employed to assess the preference of schizophrenics for varying quantities of information as presented in geometric forms. Both good premorbid paranoid and poor premorbid nonparanoid schizophrenics showed their greatest preference for the least complex shapes. The results were discussed in terms of Munsinger and Kessen's conception of a match between preference for environmental information and input-processing capacity.

Recent research has indicated that individuals show reliable preferences for different quantities of information as presented in geometric shapes (Munsinger & Kessen, 1964; Thomas, 1966). For example, Munsinger and Kessen found that Ss showed their greatest preference for figures of intermediate complexity (10 independent turns). Stimuli higher and lower in complexity were preferred less. Munsinger and Kessen have further theorized that the amount of information an individual prefers matches his information-processing capabilities. While some inconsistencies exist in this literature, it is clear that neither children nor adults prefer low-complexity shapes or, thus, low quantities of information.

In contrast, a growing body of literature has indicated that schizophrenics may prefer a low level of input. Harris (1959) reported that schizophrenics are better able than normals to tolerate the effects of sensory deprivation. McReynolds has reported that schizophrenics spent relatively less time viewing novel than familiar pictures and exhibited less stimulus seeking behavior on the Howard Maze Test (McReynolds, 1963; Sidle, Acker, & McReynolds, 1963).

These investigations, however, have not given explicit consideration to the possibility that different subgroups of schizophrenics may respond differentially to different input levels. In support of this possibility, Mehl (1967) found that among chronic nonparanoid schizophrenics only patients with poor premorbid adjustment indicated a preference for sensory deprivation. Neale and Cromwell (1968) have

proposed that among acute schizophrenics two "input processing styles" are present. Acute good premorbid and paranoid patients were hypothesized to process much information (operate at a low level of redundancy), while poor premorbid and nonparanoid patients were hypothesized to process little input (operate at a high level of redundancy).

The present study was designed to examine whether these hypothesized processing styles would be reflected in preference for complexity employing the Munsinger and Kessen paradigm. It was hypothesized that the good premorbid paranoids would show a supernormal preference for complexity and that the poor premorbid nonparanoids would show a subnormal preference for complexity.

METHOD

Nineteen good premorbid paranoid and 19 poor premorbid nonparanoid schizophrenic patients were studied. Premorbid adjustment was determined by the Phillips (1953) scale (Part I) and presence or absence of paranoid symptomatology from local psychiatric diagnosis. All patients were males between the ages of 19 and 55 and had been hospitalized for less than 3 years. Most were on standard dosages of phenothiazine medication.

The stimuli were randomly generated asymmetrical shapes. There were four stimuli at each of five complexity levels (3, 6, 10, 20, and 40 independent turns). All complexity levels were paired and any given stimulus appeared once in the series of 10 trials. The order of the stimuli within the 10 trials was determined from the procedures developed by Ross (1934). Each time a given level of complexity was paired with another level a different stimulus was employed.

The Ss were presented with the 10 pairs and on each trial were asked to indicate which of the two stimuli they preferred. This basic design was replicated four times.

¹ Requests for reprints should be sent to John M. Neale, Psychiatry Research, 110 21st Avenue, South, Nashville, Tennessee 37203.

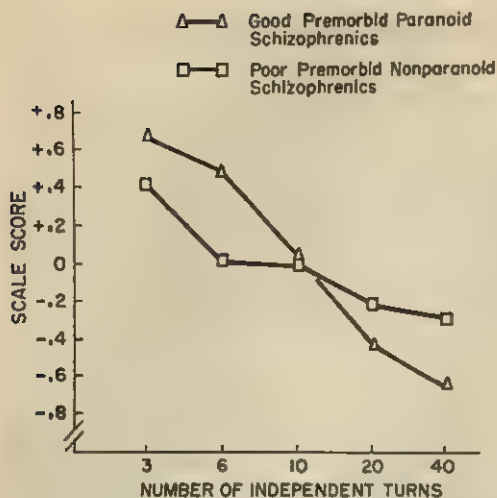


FIG. 1. Scale scores for judged preference for random shapes by the good premorbid paranoid and poor premorbid nonparanoid schizophrenics.

RESULTS AND DISCUSSION

The results were analyzed following Thurstone's Comparative Judgment Model. Scaling was accomplished under the assumptions of Classes II and III, Case C (Torgerson, 1958) and the results are presented in Figure 1. As may be seen, both the good paranoids and poor nonparanoids show their greatest preference for the least complex shapes. The number of votes for each stimulus by the two groups was analyzed by a 2 (groups) \times 5 (levels of complexity) analysis of variance. Since the assumption of homogeneity of the variance-covariance matrix could not be met, the conservative test procedure suggested by Winer (1962) was followed in evaluating the significance of the obtained F ratios. The analysis indicated a highly significant effect for levels of complexity ($F = 25.33$, $df = 1/36$, $p < .01$) and a significant Groups \times Complexity interaction ($F = 3.73$, $df = 1/36$, $p < .01$). A breakdown of this interaction indicated that the good paranoids and poor nonparanoids differed reliably from each other at Complexity Levels 6 and 40. The poors showed less preference for the 6-turn figure than the goods. For the 40-turn figure this relationship was reversed.

The results indicate that acute schizophrenics show a clear preference for stimuli of low complexity. This result has not been

found with other populations who have been tested employing this procedure (e.g., various adult populations, Munsinger & Kessen, 1964; children at various ages, Thomas, 1966). In contrast to prediction, however, there was little difference between the good paranoid and poor nonparanoid patients. In terms of Munsinger and Kessen's theory, the results may be interpreted to reflect a decreased capacity for information processing among acute schizophrenics. That is, schizophrenics prefer little information because they process information inefficiently. This reduction in the efficiency of information processing among acute schizophrenics has recently been confirmed in an investigation of span of apprehension in these patients (Neale, McIntyre, Fox, & Cromwell, 1968).

REFERENCES

- HARRIS, A. Sensory deprivation and schizophrenia. *Journal of Mental Science*, 1959, 105, 235-237.
- McREYNOLDS, P. Reactions to novel and familiar stimuli as a function of schizophrenic withdrawal. *Perceptual and Motor Skills*, 1963, 16, 847-850.
- MEHL, M. M. *The effect of brief sensory deprivation and sensory stimulation on the cognitive functioning of chronic schizophrenics*. (Doctoral dissertation, George Peabody College for Teachers) Ann Arbor, Mich.: University Microfilms, 1967. No. 67-3617.
- MUNSINGER, H., & KESSEN, W. Uncertainty, structure, and preference. *Psychological Monographs*, 1964, 78(9, Whole No. 586).
- NEALE, J. M., & CROMWELL, R. L. Size estimation in schizophrenics as a function of stimulus presentation time. *Journal of Abnormal Psychology*, 1968, 73, 44-48.
- NEALE, J. M., MCINTYRE, C. W., FOX, R., CROMWELL, R. L. Span of apprehension in acute schizophrenics. *Journal of Abnormal Psychology*, in press.
- PHILLIPS, L. Case history data and prognosis in schizophrenia. *Journal of Nervous and Mental Disease*, 1953, 117, 515-525.
- ROSS, R. T. Optimum orders for presentation of pairs in the method of paired comparisons. *Journal of Educational Psychology*, 1934, 25, 375-382.
- SIDLE, A., ACKER, M., & McREYNOLDS, P. Stimulus seeking behavior in schizophrenics and nonschizophrenics. *Perceptual and Motor Skills*, 1963, 17, 811-816.
- THOMAS, H. Preferences for random shapes: Ages six through nine years. *Child Development*, 1966, 37, 843-859.
- TORGERSOHN, W. S. *Theory and methods of scaling*. New York: Wiley, 1958.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

(Received May 22, 1968)

PATIENTS' GOALS AND THE MANIFESTATION OF PSYCHOPATHOLOGY¹

ALAN F. FONTANA²

AND

THEODORE GESSNER

*Veterans Administration Hospital
West Haven, Connecticut*

University of Maryland

Three groups of psychiatric patients in a veterans general hospital were tested under research, transfer, and discharge instructional sets. Differences in self-reported deviance among the groups are interpreted within an impression management framework for the information they reveal about patients' motivational orientations toward transfer and discharge as hospitalization outcomes. Patients were of greatest accord in wanting to avoid transfer. Orientation toward discharge varied among patient subgroups. Shifts in the level of presented deviance were similar for all dimensions but occurred significantly ($p < .05$) only for psychopathology. Differential levels of presented deviance among subgroups are interpreted in terms of corresponding differences in judged appropriateness for patienthood.

The assertion that "psychopathological" behavior can be adequately understood as one way of attempting to exercise control over current interpersonal interactions (e.g., Haley, 1963; Szasz, 1964) is a position with numerous implications. One of these is that the behavior of mental patients and the "mentally ill" can be analyzed in terms of self-presentation or impression management just as fruitfully as the behavior of nonpatients or "normals" can. The transactional nature of psychological testing provides a good example of the applicability of this view. Psychological tests are usually given in conjunction with a forthcoming decision to be made about a patient. If psychological testing is considered to be a two-way communication situation as Rotter (1960) and others have contended, the patient's perception of the nature of the decision must be considered an important determinant of his responses and not simply a biasing factor. The patient receives certain benefits from the hospital or clinic, and his presentation on psychological tests is one of the major means available to him for influencing decisions in the direction of his goals.

When the communication approach to psychological testing is applied, inferences about the types and strengths of patients' goals can be made on the basis of response changes when the avowed purpose of the testing changes. Several investigators have reported that patients manifest differing amounts of pathology depending upon their understanding of the reasons for testing and the possible relevance of the results to the attainment of their goals (Braginsky & Braginsky, 1967; Braginsky, Grosse, & Ring, 1966; Wilcox & Krasnoff, 1967; Young, 1965). In a similar vein, Fontana and Klein (1968) categorized schizophrenic patients as healthy or sick presenters on the basis of their self-reports. When tested on reaction time, the deficit manifested by patients varied from none to a substantial amount and was found to be directly related to the adequacy of self-presentation and the characteristics of the hospital environment. Explanation of the results included the assertion that patients in treatment buildings were motivated to maintain the status quo; that is, they neither wanted to be transferred to custodial buildings nor to be discharged. The present study was designed to test the generality of this assertion by varying the meaning of the testing conditions under which patients in a different treatment hospital completed several self-report scales. The expectation was that patients would attempt to create the impres-

¹The authors wish to thank Mary McIndoo, Jonathan Cummings, and the staff of the Veterans Administration Hospital, Washington, D. C., for their support. Appreciation is also due Edward Klein and Dorothea Braginsky for their critical reading of the manuscript.

²Also at Yale University. Requests for reprints should be sent to Alan F. Fontana, Veterans Administration Hospital, West Haven, Connecticut 06516.

sion that they were too healthy to warrant transfer but too sick to deserve discharge.

Psychopathology and adequacy-responsibility are two dimensions considered to be of central importance to the transactions between patients and staff. Psychopathology should be expected to occupy a central position because hospitals are established for the treatment of pathology as their reason for being. The strong emphasis in clinical thinking on patients' need for self-esteem and the acceptance of responsibility suggests adequacy-responsibility as an important dimension also. In the self-presentation view of psychopathology both psychotics and nonpsychotics are considered to be equally aware of the demands and norms of their environment. It is recognized that the mode of impression management is different for the groups, but it is expected that both groups attempt to manage their impressions. Three testing conditions were employed in which testing was explained in terms of research interests. Further information was provided in two conditions, one in which patients were told that the results might have personal relevance to them regarding transfer to another hospital, and the second in which they were told that the results might have personal relevance to them regarding discharge from the hospital.

The major hypotheses can be stated as follows:

1. Compared to patients in the research-only condition, patients in the transfer-relevant condition would attribute less psychopathology to themselves, and patients in the discharge-relevant condition would ascribe more pathology to themselves.
2. Compared to patients in the research-only condition, patients in the transfer-relevant condition would describe themselves as more adequate-responsible, and patients in the discharge-relevant condition would report that they were less adequate-responsible.
3. Comparable effects would be found for patients diagnosed as psychotic and non-psychotic.

METHOD

Subjects and Setting

Seventy-five male psychiatric patients from a Veterans Administration general medical and surgical

hospital comprised the *S* sample. Forty were diagnosed nonpsychotic (68% assigned diagnosis of anxiety reaction, depressive reaction, or alcoholism) and 35 psychotic (89% schizophrenic). The average patient was 41.4 years of age ($SD = 8.8$), had 11.4 years of education ($SD = 3.1$), and had been hospitalized currently for 5.4 months ($SD = 3.8$). Sixteen of the *Ss* were Negro and 59 were Caucasian. Assignment of *Ss* to testing conditions was balanced for diagnosis, race, and ward of residence. The *Ss* in the three conditions did not differ significantly in age, education, or length of hospitalization. Excluded from the sample were patients considered to be organically impaired, those 60 years of age or older, and persons currently hospitalized less than 1 month.

The hospital is a new, attractive building located in the center of a metropolitan area. All wards are open, extensive facilities are available, and an active treatment program is in force. A psychiatric transfer for most patients would entail going to older hospitals, located in rural areas, where less extensive facilities and treatment are available. Klein and Fontana (in press) found that patients evaluate hospitals of the former type much more positively than hospitals of the latter type. It is reasonable to conclude that patients would wish to avoid being transferred from a treatment hospital.

Testing Conditions and Instruments

The *Ss* were tested in groups and were told that the testing was being conducted for research purposes. For the 27 *Ss* in the research condition this was their only information. The transfer group ($N = 24$) was additionally informed,

I should add that this is mainly a research study but that it may have personal relevance to you since one of the purposes of this study is to determine which patients have received maximum benefit from this hospital and should be transferred to _____ (a large psychiatric hospital).

The discharge group ($N = 24$) was given this information with the exception that discharge was substituted for transfer. The phrase "maximum benefit" was employed because it is the language, familiar to both staff and patients, used to refer to a staff judgment that further treatment is unlikely to be productive and that the patient's association with the hospital should be terminated. The research group was tested 1 week before the transfer and discharge groups, which were tested at the same time. In this way, there was no possibility that the discharge or transfer instructional set might influence any testing condition other than the one intended. Psychopathology was considered to be the most prominent dimension for manipulation. A scale assessing the attribution of psychopathology to oneself (SD_{1a}) reported by Fontana, Klein, Lewis, and Levine (1968) and a scale for predicting chronicity (*C*) developed by Anker (1961), were used as indexes of psychopathology.

TABLE 1

MEANS AND ANALYSES OF VARIANCE OF THE RESPONSES TO TESTING CONDITIONS
BY PSYCHOTICS AND NONPSYCHOTICS

Variable	Diagnosis	Conditions			Analyses of variance		
		Research	Transfer	Discharge	Conditions	Diagnosis	Interaction
SD_{18}	Psychotic	11.31	13.67	12.23	*	**	
	Nonpsychotic	9.00	12.07	9.00			
C	Psychotic	6.85	5.00	6.08	*		
	Nonpsychotic	7.43	3.93	5.91			
K'	Psychotic	9.31	9.22	10.54		**	
	Nonpsychotic	5.79	8.67	6.18			
AR	Psychotic	9.77	12.56	11.31		**	
	Nonpsychotic	6.93	9.20	8.45			
I-E	Psychotic	10.38	9.33	9.54			
	Nonpsychotic	12.29	9.53	9.91			
ΣT	Psychotic	38.92	40.89	38.08	*		*
	Nonpsychotic	46.36	39.47	37.82			

Note.—For psychotics in the research group, $N = 13$; in the transfer group, $N = 9$; in the discharge group, $N = 13$. For nonpsychotics in the research group, $N = 14$; in the transfer group, $N = 15$; in the discharge group, $N = 11$.

* $p < .05$.

** $p < .01$.

In order to resolve an interpretive ambiguity surrounding pathology scales such as SD_{18} and C , Welsh's "purification" of the MMPI denial scale (K') (Dahlstrom & Welsh, 1960) was included. K' consists of the original K scale minus those items also included in the MMPI pathology scales. Thus K' can be taken as a measure of the denial of undesirable but nonpathological statements. If Ss respond to K' items in the same way and to the same extent as they do to SD_{18} and C items, it would suggest that self-presentation is largely achieved through a general denial of undesirable characteristics. If, however, Ss respond somewhat differently to K' than to SD_{18} and C , it would suggest that the pathology content of the items is a crucial variable.

The second dimension thought to be relevant to impression management is adequacy-responsibility (AR). A 20-item, counterbalanced, true-false scale was devised which included items such as: "I can successfully accomplish most things I set out to do"; "I am someone who people can depend on"; and "I think that a person with an emotional problem is still accountable for his actions." A second scale used to assess this dimension was the Internal-External Locus of Control Scale (I-E) developed by Rotter (1966).

Finally the sum of the "true" responses (ΣT) to the counterbalanced scales was calculated. This served as a measure of the extent to which Ss adopted an uncritical, uninvolved orientation to the items as opposed to a critical involvement in the content of the items.³

³ The split-half reliabilities, corrected by the Spearman-Brown formula, for these measures are as follows: SD_{18} (.84), C (.80), K' (.84), AR (.80), I-E (.77), and ΣT (.53).

RESULTS

A 3 (testing conditions) \times 2 (diagnoses) analysis of variance was performed for each scale and the sum of the "true" responses. The means and results of the analyses of variance are presented in Table 1. High scores on SD_{18} , AR, and K' indicate a healthy, adequate, denying orientation, respectively; whereas high scores on C and I-E indicate an orientation toward greater chronicity and external locus of control, respectively.

Inspection of the means reveals that there is a consistent pattern of response to all scales across the testing conditions. As predicted, Ss responded in the least deviant manner in the transfer condition. However, this effect was significant only for measures of the pathology dimension, SD_{18} and C . The Ss did not respond in the most deviant manner in the discharge condition as expected; rather the means typically fell between those for the research and transfer conditions. Psychotics consistently presented themselves in a less deviant light than nonpsychotics, significantly so on SD_{18} , K' , and AR. As hypothesized, the psychotics' pattern of responses to the testing conditions was parallel to that of nonpsychotics in every case. ΣT was the only variable to yield a significant interaction between

groups and conditions; nonpsychotics endorsed more items as true than did psychotics in the research condition.

An exploratory analysis was conducted involving patients' service connection and disability status. This variable was thought to be relevant to patients' hospitalization goals since service-connected patients have priority in claiming treatment resources and are the only ones automatically eligible for psychiatric disability payments. Three groups were established: nonservice-connected, service-connected with partial disability, and service-connected with full disability. To the extent that considerations regarding disability ratings and payments affect hospitalization goals, nonservice-connected patients could be expected to be affected least by differences in the testing conditions and partial disability patients to be affected most, particularly in the comparison between research and discharge conditions. All service-connected pa-

tients receive disability payments at the rate of 100% while they are in the hospital, providing they have been hospitalized more than 21 days. Upon discharge, payments revert to the level of the prehospitalization rating. Discharge would mean no decline in payments to the fully disabled, but a decrease for the partially disabled. Thus, to the extent that a reduction in payments might influence hospitalization goals, the partially disabled could be expected to be most resistant to discharge and to present themselves with greatest deviance in the discharge condition. Patients in the three groups were comparable in terms of diagnoses and background variables except for prior hospitalization history; men with a 100% disability rating (prior to current hospitalization) had been hospitalized for psychiatric problems significantly more often ($F = 6.57$, $df = 2/72$, $p < .01$) than patients in the other groups. The means and results

TABLE 2
MEANS AND ANALYSES OF VARIANCE OF THE RESPONSES TO TESTING CONDITIONS
BY NONSERVICE-CONNECTED, AND PARTIALLY DISABLED AND FULLY
DISABLED SERVICE-CONNECTED PATIENTS

Variable	Service connection status	Conditions			Analyses of variance		
		Research	Transfer	Discharge	Conditions	Service connection	Interaction
SD_{18}	Nonservice connected	11.40	11.78	11.50			
	Partially disabled	8.43	13.89	11.83	**	*	
	Fully disabled	8.60	12.17	6.50			
C	Nonservice connected	6.07	5.22	6.36			
	Partially disabled	8.71	2.78	3.67	**	*	
	Fully disabled	8.20	5.33	8.25			
K'	Nonservice connected	8.53	8.67	9.64			
	Partially disabled	5.57	9.22	8.50			
	Fully disabled	7.00	8.67	4.75			
AR	Nonservice connected	9.53	10.44	11.00			
	Partially disabled	7.29	11.33	11.00		**	
	Fully disabled	6.00	9.17	5.00			
$I-E$	Nonservice connected	10.53	9.78	10.43			
	Partially disabled	13.00	8.33	7.83			
	Fully disabled	11.60	10.67	10.00			
ΣT	Nonservice connected	41.00	38.89	38.43			
	Partially disabled	45.57	41.67	37.50	**		
	Fully disabled	44.20	39.17	37.00			

Note.—For the nonservice-connected group in the research condition, $N = 14$; in the transfer condition, $N = 15$; in the discharge condition, $N = 9$; in the condition, $N = 9$; in the discharge condition, $N = 6$. For the fully disabled service-connected patients in the research condition, $N = 7$; in the transfer condition, $N = 5$; in the discharge condition, $N = 6$; in the discharge condition, $N = 4$.

* $p < .10$.

** $p < .05$.

of the analysis of variance can be found in Table 2.

Inspection of the means indicates that non-service-connected patients were much less affected by the testing conditions than were service-connected patients. The data also reveal that the service-connected groups presented themselves similarly in the research condition and with least deviance in the transfer condition, but differed markedly in their responses to the discharge condition. Compared to the research condition, full disability patients attributed the same or greater deviance to themselves in the discharge condition. Contrary to expectations, partial disability veterans described themselves much less deviantly in the discharge condition than in the research condition.

DISCUSSION

The results have strongly supported some of the conceptions about patients' self-presentation and have indicated that modifications are needed in others. The prediction that patients would not want to be transferred to a less desirable hospital was clearly borne out. Under classifications of both diagnosis and service connection, patients in the transfer condition presented themselves less pathologically than those in the research condition. Patients' orientations toward discharge were less clear. When Ss were classified by diagnosis, the level of reported pathology in the discharge condition was greater than in the transfer condition and usually similar to the research condition. In addition, the variability of scores was consistently (though not significantly) found to be largest in the discharge condition and smallest in the transfer condition. This suggested that there was a greater difference of opinion among patients concerning the desirability of discharge compared to transfer, and that patients were either indifferent to discharge or divided, with one subgroup favoring and one opposing discharge. The classification according to service connection helped to clarify this ambiguity and indicated that all three orientations were represented.

Non-service-connected patients varied very little in their presentations, demonstrating an indifference to the added information in the

transfer and discharge conditions. Full and partial disability Ss differed markedly in their presentations, the fully disabled manifesting more deviance than the partially disabled. The difference was most pronounced in the discharge condition. The fully disabled manifested a pattern consistent with avoiding discharge whereas the partially disabled produced a pattern consistent with obtaining discharge. The latter finding is particularly unexpected because there were plausible reasons for suspecting that the partially disabled would be most desirous of avoiding discharge. Certainly this finding runs counter to the widespread impression that potential reduction of disability payments encourages adoption of the sick role and the desire for hospitalization. These data suggest that analyses and explanations based solely in terms of monetary payments are oversimplifications and that the effects of disability payments result from a complex interaction of factors little understood at present. A similar conclusion was drawn by Fisher (1965), who found that disability payments did not compensate for low occupational adjustment in any simple and direct way.

The fact that the fully disabled confirmed the initial hypothesis and the partially disabled did not may be related to the greater number of prior hospitalizations of the former group. It is not surprising that the orientation of staying in the hospital was found for those patients who had a history of repeated hospitalization. A chronicity difference might also account for the authors' general overstatement of patients' desire to avoid discharge. Patients in the present sample were currently hospitalized an average of only 5.4 months compared to 22.9 months for the Veterans Administration sample investigated by Fontana et al. (1968). Thus patients in the latter sample may have created the stronger and more pervasive impression of wanting to stay in the hospital concomitant with their greater chronicity.

Psychotics consistently presented themselves less deviantly than did nonpsychotics. The fact that the difference was most evident in the research condition indicates that the groups approached testing with different orientations. This conclusion is supported further

by the difference in the sum of the "true" responses between diagnostic groups. Nonpsychotics in the research condition chose the "true" alternative much more often than did the psychotics or the nonpsychotics in the other conditions. Thus the nonpsychotics responded less to the content of the items. Psychotics maintained close attention to item content throughout the three conditions. A similar difference in the research condition can be seen for the service connection classification. Nonservice-connected veterans presented themselves less deviantly than did service-connected patients. Although the diagnostic and service connection classifications are independent ($\chi^2 = 1.96$, $df = 2$, $p > .25$), they both possess an internal classification of greater and lesser appropriateness for admission and retention.

The fact that this was a treatment hospital and that psychotics are generally considered to be the poorer candidates for individual and group therapy could be expected to exert a selective influence in favor of the healthier appearing psychotics. Similarly, a greater weight may have been placed on healthier self-presentations by nonservice-connected veterans because of their lesser priority in claiming hospital resources. Conversely, service-connected veterans because of their legal status, and nonpsychotic patients as a result of their symptomatology, could be expected to be a less selected sample in terms of their self-presentations by virtue of their greater suitability for patienthood.

With the exception of the "true" responses, nonpsychotics and psychotics responded to the testing conditions in a parallel manner. Both groups revealed an awareness of the differential meaning of the testing conditions and an interest in varying the level of deviance presented.

Impression management mainly involved scales with pathology content. Scales measuring adequacy-responsibility were only weakly affected by the testing conditions, albeit in the same direction as measures of psychopathology. This is surprising since so much of therapeutic endeavor is conceptualized in terms of increasing a patient's sense of personal responsibility. However, patient responsibility is often incompatible with the tradi-

tional patient role, and it is very possible that double messages are often given to hospitalized patients (Wilensky & Herz, 1965). Explicit expectations of responsibility are generally expressed through words, while covert expectations against responsibility are expressed through actions. The resulting ambiguity could be expected to make patients unsure of how to present themselves on this dimension.

The greater sensitivity to the testing conditions of the pathology scales as compared to the denial scale helps to resolve a question of interpretation of the former. Although K' correlated .75 and $-.52$ with SD_{18} and C , respectively, the shift in K' scores across conditions was not statistically significant. It seems reasonable to conclude that although a general defensiveness or social desirability orientation was operating, pathology content as focused avenue of self-presentation was important over and above such an orientation.

REFERENCES

- ANKER, J. M. Chronicity of neuropsychiatric hospitalization: A predictive scale. *Journal of Consulting Psychology*, 1961, 25, 425-432.
- BRAGINSKY, B. M., & BRAGINSKY, D. D. Schizophrenic patients in the psychiatric interview: An experimental study of their effectiveness at manipulation. *Journal of Consulting Psychology*, 1967, 31, 546-551.
- BRAGINSKY, B. M., GROSSE, M., & RING, K. Controlling outcomes through impression-management: An experimental study of the manipulative tactics of mental patients. *Journal of Consulting Psychology*, 1966, 30, 295-300.
- DAHLSTROM, W. G., & WELSH, G. S. *An MMPI handbook*. Minneapolis: University of Minnesota Press, 1960.
- FISHER, G. J. Socioeconomic factors in outcome of released mental patients: Influence of type of placement, occupational adjustment, compensation and type of hospital. *Journal of Health and Human Behavior*, 1965, 6, 105-110.
- FONTANA, A. F., & KLEIN, E. B. Self-presentation and the schizophrenic "deficit." *Journal of Consulting and Clinical Psychology*, 1968, 32, 250-256.
- FONTANA, A. F., KLEIN, E. B., LEWIS, E., & LEVINE, L. Presentation of self in mental illness. *Journal of Consulting and Clinical Psychology*, 1968, 32, 110-119.
- HALEY, J. *Strategies of psychotherapy*. New York: Grune & Stratton, 1963.
- KLEIN, E., & FONTANA, A. F. Understanding the schizophrenic. *International Journal of Social Psychology*, in press.

- ROTTER, J. B. Some modifications of social learning theory for the prediction of goal directed behavior from testing procedures. *Psychological Review*, 1960, 67, 301-316.
- ROTTER, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 1966, 80(1, Whole No. 609).
- SZASZ, T. S. *The myth of mental illness*. New York: Hoeber-Harper, Harper & Row, 1964.
- WILCOX, R., & KRASNOFF, A. Influence of test-taking attitudes on personality inventory scores. *Journal of Consulting Psychology*, 1967, 31, 188-194.
- WILENSKY, H., & HERZ, M. I. The social role of the psychiatric patient in a general hospital. *Proceedings of the 73rd annual convention of the American Psychological Association*, 1965. Washington, D. C.: American Psychological Association, 1965.
- YOUNG, R. C. Effects of differential instructions on the Minnesota Multiphasic Personality Inventories of state hospital patients. *Proceedings of the 73rd annual convention of the American Psychological Association*, 1965. Washington, D. C.: American Psychological Association, 1965.

(Received May 27, 1968)

(Continued from page 239)

- Prospective Span as a Cognitive Ability: Louis S. Dickstein*: Department of Psychology, The City College of the City University of New York, New York 10031.
- Elimination of Sleep Walking in a Seven-Year-Old Boy: Paul W. Clement*: Graduate School of Psychology, Fuller Theological Seminary, 177 North Madison Avenue, Pasadena, California 91101.
- Bias in the Assessment of Psychotherapy: John C. Steinhelber*: Veterans Administration Hospital, Forty-Second Avenue and Clement Street, San Francisco, California 94121.
- Self-Other Orientations of Institutionalized Behavior-Problem Adolescents: Barbara H. Long,* Robert C. Ziller, and John Bankes: Department of Psychology, Goucher College, Baltimore, Maryland 21204.

* Asterisk indicates author for whom address is supplied.

RELATIONSHIP OF SELECTED PSYCHOSOCIAL VARIABLES TO PROGNOSTIC JUDGMENTS

JAMES E. COWDEN¹ AND ASHER R. PACTH

Wisconsin Department of Health and Social Services

The way in which clinicians use selected cue variables in making prognostic judgments of the postrelease adjustment of institutionalized delinquents was assessed in this study. Cue variables (concepts commonly used in clinicians' reports, typically with some prognostic significance) were correlated with prognostic judgments of postrelease adjustment and also with actual ratings of postrelease adjustment. The clinicians were reasonably accurate, as a group, in estimating the prognostic significance of the cue variables. Cue variables more readily inferred from behavior (institutional adjustment, hostility, home environment, and maturity) were related most consistently and accurately to postrelease prognosis, while those less easily inferred from behavior (anxiety, dependency, and guilt) showed a less clear-cut relationship. More communication among clinicians, particularly in their usage of concepts describing intrapsychic processes, was thus recommended.

Isolating the factors involved in making clinical judgments has long been a problem for the clinician. This problem is further complicated when the clinician is asked to make diagnostic assessments in a correctional setting. Most clinical instruments have been standardized on noncorrectional populations. Inmates tend to be rather defensive and uncommunicative, and there is typically considerable social class distance between the clinicians and the inmates. These factors in combination have served to heighten the difficulty of making meaningful clinical predictions (Harari, 1964; Jacks, 1962; Prentice, 1964). As a result, false inferences are drawn and inmates are frequently stereotyped as being sick and/or victimized by their environment (Giedt, 1955; Lee & Tucker, 1962; Kostlan, 1954). More effective prediction can best be achieved through a systematic analysis of the relationships of the factors which underlie clinical judgment.

This study examined the accuracy and consistency with which clinicians utilized selected cue variables in making prognostic judgments relating to the postrelease adjustment of delinquent boys.

METHOD

Four experienced staff members (three psychologists and a psychiatrist) at the Wisconsin School for

¹ Requests for reprints should be sent to James E. Cowden, 1 West Wilson Street, P. O. Box 669, Madison, Wisconsin 53701.

Boys, Wales, were surveyed to determine which factors they considered to be most predictive of the postrelease adjustment of delinquent boys. From among those typically included in reports the nine factors selected for this study were: hostility, anxiety, dependency, overall magnitude of personality disturbance, social delinquent features (degree to which involvements with delinquent peers contributed to their delinquent behavior), institutional adjustment, level of maturity, and adequacy of their home environment. These factors were labeled "cue variables" and a 7-point rating scale was constructed for each. In addition, a 7-point "global prognostic" rating scale was developed for clinicians to make predictions concerning the postrelease adjustment of each boy in the sample. Finally, a 7-point rating scale was constructed to measure the actual postrelease adjustment of Ss for a period of up to 1 year following their release from the institution.

Ratings were made on each of the nine cue variables, the global prognostic scale of postrelease adjustment, and the rating scale of actual postrelease adjustment on 152 delinquent boys who had been institutionalized as delinquents at the Wisconsin School for Boys at Wales.

Each clinician rated approximately the same number of boys and no boy was rated by more than one clinician. Interjudge reliabilities on these scales had been assessed previously and are reported in the results section. To avoid bias, clinicians were not aware that the study involved correlating ratings on each of the cue variables with global prognostic ratings and actual ratings of postrelease adjustment.

To assess the predictive value of each cue variable, the cue variable ratings of each of the four raters (separately and combined) were correlated both with their global prognostic ratings of postrelease adjustment and with ratings of actual postrelease adjustment of Ss made independently by other raters. The results of a previous study (Cowden & Pacht, 1967)

TABLE 1
CORRELATIONS OF CUE VARIABLES WITH GLOBAL PROGNOSTIC RATINGS OF POSTRELEASE
ADJUSTMENT AND WITH RATINGS OF THE ACTUAL POSTRELEASE
ADJUSTMENT OF 152 DELINQUENT BOYS

Cue variables*	Global prognostic ratings of postrelease adjustment					Ratings of actual post-release adjustment
	Separate raters				Combined raters	
	JC	TB	WP	GM		
Institutional adjustment	.65	.78	.75	.71	.65	.39
Hostility	-.40	-.11	-.53	-.16	-.36	-.15
Home environment	.31	.26	.38	.17	.28	.17
Immaturity	-.35	-.07	-.21	-.34	-.24	-.16
Personality disturbance	-.23	-.51	.25	-.43	-.32	-.11
Guilt	.39	-.29	.63	.06	.29	.18
Social delinquent features	-.23	.09	-.41	-.10	-.24	-.14
Dependency	.35	-.19	.37	.08	.16	-.15
Anxiety	-.10	-.12	.52	.00	-.02	.02

Note.—Global prognostic ratings of postrelease adjustment were made by psychologists from reports prepared by clinicians (psychologists or psychiatrists) shortly after admission to the institution. Ratings of actual postrelease adjustment were made from parole agents' chronological reports of their adjustment up to 1 year following their release from the institution.

* Ranked in order of the consistency of ratings and overall magnitude (either positive or negative) of the correlation coefficients.

indicated that clinicians' global prognostic ratings are among the best predictors of the subsequent behavior of institutionalized delinquents. In that study, it was found that clinicians' global prognostic ratings showed a moderately high correlation ($r = .51$) with rated postrelease adjustment. It was, therefore, assumed that the correlation between a clinician's ratings of a boy on a given cue variable and his global prognostic ratings of the same boy would reflect the importance of this cue variable as a source of information for the clinicians in making judgments as to the boy's probable postrelease adjustment. Correlations between cue variables and prognostic ratings for each clinician separately were then compared to determine degree of consistency among clinicians in utilizing these cue variables as sources of information in making judgments about the postrelease adjustment of delinquent boys.

The foregoing procedures permitted the following comparisons to be made: First, the consistency with which clinicians utilized the cue variables in making prognostic judgments was determined by assessing the degree of relationship between ratings of each cue variable and the global prognostic ratings for each clinician in turn. Second, the relative accuracy with which clinicians as a group used each of the cue variables as predictors of postrelease adjustment was determined by an assessment of the relationship between the cue variables, global prognostic ratings, and ratings of actual postrelease adjustment for all clinicians combined.

RESULTS

Interrater reliabilities of each of the nine cue variables and the global prognostic rating

variable were assessed through having two psychologists make independent ratings of 35 delinquent boys on each of these variables. The reliability coefficients for these variables were as follows: institutional adjustment, .85; maturity, .83; hostility, .78; social delinquent features, .76; personality disturbance, .73; home environment, .71; guilt, .66; anxiety, .56; dependency, .52; global prognosis, .79. These variables can thus be rated by experienced clinicians with at least moderate reliability.

The cue variables were then correlated with ratings of the actual postrelease adjustment of the boys, and, second, with the global prognostic ratings of their postrelease adjustment. The results of this analysis are shown in Table 1.

These results indicate that the cue variable, institutional adjustment, showed the highest correlation both with the global prognostic and actual ratings of postrelease adjustment. In addition, the raters showed a high degree of consistency in utilizing this cue variable in making their global prognostic ratings. In similar fashion, the variables of hostility, home environment, and maturity (in that order) were also, to a lesser extent, used consistently by clinicians in making their

global prognostic ratings of postrelease adjustment. In particular, the results suggested that delinquents with a better than average institutional adjustment, who came from a relatively good home environment, and who showed relatively little hostility and above-average maturity, would be rated by most clinicians as having a positive prognosis. The variables of dependency and anxiety, on the other hand, were not consistently correlated with the clinicians' prognostic judgments.

A comparison of the correlations between the cue variables and global prognostic ratings, and between the cue variables and actual postrelease adjustment for the sample of delinquent boys, demonstrated reasonably good agreement in the direction and magnitudes of the correlation coefficients. The magnitudes of the correlations between the cue variables and global prognostic ratings were, however, somewhat higher than those between cue variables and actual postrelease adjustment, which suggests that clinicians tended to give the cue variables more "weight" as predictors than was warranted. The relative weightings given to each of the cue variables appeared realistic, however, as indicated by the correlation (ϕ coefficient) of .74 between the relative rankings of the magnitudes of the correlation coefficients between the cue variables and global prognostic ratings, and between the cue variables and ratings of actual postrelease adjustment. It seems clear that experienced clinicians are moderately accurate not only in predicting the postrelease adjustment of institutionalized delinquents, but also in the degree to which they utilize various cue variables in making such predictions.

DISCUSSION

The results of this study suggest that under certain conditions clinicians show considerable inconsistencies as to which variables are most clearly related to the postrelease adjustment of delinquent boys. Their degree of consistency depended heavily upon the nature of the concepts being rated. It was found, for example, that the variables of institutional

adjustment, hostility, home environment, and maturity were most consistently related to postrelease prognosis. These results suggest that those concepts most readily inferred from behavior can be more clearly and consistently related to prognostic judgments by clinicians. By contrast, the variables of guilt, anxiety, and dependency were least consistently related to prognostic judgments. These results are particularly meaningful since these latter concepts are so frequently mentioned in clinicians' reports, typically with some prognostic implications. The lack of consistency among clinicians in their perceptions of the relationships between these concepts and predicted postrelease adjustment appears to be a consequence, at least in part, of the relative ambiguity of these concepts.

In general, the results of this study suggest that (a) clinicians should be cautious in assuming a significant consensus as to the prognostic significance of many concepts frequently found in clinical reports; (b) more communication among clinicians should be encouraged with the goal of achieving greater mutual understanding and more consistent usage of certain key concepts, particularly those describing intrapsychic processes.

REFERENCES

- COWDEN, J., & PACT, A. Predicting institutional and postrelease adjustment of institutionalized delinquents. *Journal of Consulting Psychology*, 1967, 31, 377-381.
- GIEDT, F. H. Comparison of visual, content, and auditory cues in interviewing. *Journal of Consulting Psychology*, 1955, 19, 407-416.
- HARARI, C. Class bias in psychodiagnosis of delinquents. *Crime and Delinquency*, 1964, 10, 145-151.
- JACKS, I. Psychological testing in institutional rehabilitation centers—a gentle critique. *Crime and Delinquency*, 1962, 8, 34-39.
- LEE, J. C., & TUCKER, R. B. An investigation of clinical judgment: A study in method. *Journal of Abnormal and Social Psychology*, 1962, 64, 272-280.
- KOSTLAN, A. A method for the empirical study of psychodiagnosis. *Journal of Consulting Psychology*, 1954, 18, 83-88.
- PRENTICE, N. M. Psychological testing in the correctional institution—another viewpoint. *Crime and Delinquency*, 1964, 10, 263-268.

(Received June 7, 1968)

BRIEF REPORTS

COMPARATIVE STUDY OF THE WECHSLER PRESCHOOL AND PRIMARY SCALE OF INTELLIGENCE AND THE STANFORD-BINET INTELLIGENCE SCALE, FORM L-M, AMONG CULTURALLY DEPRIVED CHILDREN¹

ALLAN BARCLAY AND ALLAN C. YATER

St. Louis University

This study investigated the relationship between the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, 1967) and the Stanford-Binet Intelligence Scale, Form L-M (S-B; Terman & Merrill, 1960), among a sample of culturally disadvantaged children in order to afford some estimate of the relative similarities or differences that might exist between these two instruments.

The Ss were 50 culturally disadvantaged children who had been participating in a Head Start program for an average period of 8 months. The sample was drawn from neighboring centers in an attempt to minimize differences in residential environments. The children ranged in age from 60 through 71 months ($M = 63.84$, $SD = 2.44$). Two female examiners administered the tests, and each child was tested individually in his respective Head Start center. Ten subtests of WPPSI were administered in the order suggested by Wechsler (1967). The omitted subtests were Sentences (Supplementary) and Animal House Retest.

A counterbalanced order of presentation of the WPPSI and the S-B was employed, with a mean interval between test administrations of 21 days. Twenty-five Ss received the WPPSI first; the remaining Ss completed the S-B first. Each group contained 13 males, 12 females, 11 Caucasians, and 14 Negroes.

The means and standard deviations for the WPPSI Verbal, Performance, and Full-Scale IQs were 91.72 and 10.97, 96.32 and 12.09, 93.38 and

11.57, respectively. The mean Performance IQ was significantly higher than the mean Verbal IQ ($t = 3.34$, $df = 49$, $p < .01$) and the mean Full-Scale IQ ($t = 4.09$, $df = 49$, $p < .001$), while the mean Full-Scale IQ was significantly higher than the mean Verbal IQ ($t = 2.39$, $df = 49$, $p < .05$).

The mean S-B Full-Scale IQ was 100.96 with a standard deviation of 12.46. The S-B IQ mean was significantly higher than the mean WPPSI Verbal IQ ($t = 7.39$, $df = 49$, $p < .001$) as well as the mean WPPSI Performance IQ ($t = 3.69$, $df = 49$, $p < .001$) and the mean WPPSI Full-Scale IQ ($t = 7.13$, $df = 49$, $p < .001$).

A 2×2 Latin-square design also compared the mean WPPSI Full-Scale IQ with the mean S-B IQ. It was again found that the mean S-B Full-Scale IQ was significantly higher than the mean WPPSI Full-Scale IQ ($F = 51.73$, $df = 1/48$, $p < .001$). However, order of administration (i.e., WPPSI-S-B vs. S-B-WPPSI) had no significant effect on IQ scores ($F < 1.00$, $df = 1/48$).

Further comparisons were made between males and females, between Negroes and Caucasians, and between children 60-63 months and 64-71 months of age. No significant differences due to sex, race, or age were found on any of the four IQ estimates.

Correlation coefficients for the S-B IQ and the WPPSI Verbal, Performance, and Full-Scale IQs were .73, .74, and .81, respectively.

Results were interpreted as suggesting that the WPPSI and the S-B were not comparable instruments for the culturally disadvantaged children in this study, and that the WPPSI seemed to be experienced as the more difficult test.

REFERENCES

- TERMAN, L. M., & MERRILL, M. A. *Stanford-Binet Intelligence Scale—Form L-M: Manual*. Boston: Houghton Mifflin, 1960.
WECHSLER, D. *Wechsler Preschool and Primary Scale of Intelligence*. New York: Psychological Corporation, 1967.

(Received August 12, 1968)

¹ An extended report of this study may be obtained without charge from Allan Barclay, Department of Psychology, St. Louis University, St. Louis, Missouri 63103, or for a fee from the American Science Information Service. Order NAPS Document NAPS-00272 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001; remitting \$1.00 for microfiche and \$3.00 for photocopies. Requests for reprints should be sent to Allan Barclay at the above address.

REVISED SUICIDE POTENTIAL SCALE¹

R. W. MISKIMINS AND LOWELL T. WILSON

Colorado State University

A problem of extreme interest for all psychiatric institutions is that of suicide prevention. An investigation by Miskimins, DeCook, Wilson, and Maley (1967) identified readily obtained clues in the demography and personality of institutionalized mental patients who had committed suicide. These empirically derived items were cast together in the form of a checklist and entitled the Suicide Potential Scale (SPS). The SPS construction was completed in 1965 and staff-selected potential suicides provided the only means for "cross-validation" of the instrument. With the passage of 3 years, there became available another sample of suicides, and hence, the present study was undertaken.

For the original research, study groups were established by matching 16 psychiatric patients who had committed suicide with nonsuicidal patients on the variables of sex, age, marital status, and diagnosis. The two groups were then compared on routine demographic and psychiatric information that was collected prior to the incidence of self-destructive behavior. The SPS was constructed from 23 items which were found to discriminate the two groups, plus an additional three based upon variables employed in the matching.

The Ss for the present study were drawn from the patient population of the Fort Logan Mental Health Center, Denver, Colorado. The experimental groups consisted of 15 patients who had committed suicide during the period from June, 1965 to June, 1968. A nonsuicidal control group, totaling 30, was selected randomly from the re-

mainder of the hospital population. These two patient groups were scored on the original SPS. The suicide group mean score was 32.60, as compared to 22.37 for the control group ($t = 6.25$, $p < .001$). An item analysis was undertaken to cull out the least discriminative items, both to increase the scale's precision and to decrease its length. This procedure suggested the elimination of 10 low-differentiation items; the remaining 16 were used to construct the Revised Suicide Potential Scale (RSPS). The 16 RSPS items were based on the following characteristics: sex and age, diagnosis, times admitted, marital status, education, preoccupation, slowing of thought, language use, anger, depression, apathy, inappropriate behavior, social pattern against, impaired effectiveness, external precipitating stress, and danger to self.

All six study groups, from the original and present research, were scored on the RSPS. The mean scores for these samples (and the Ns on which they were based) were as follows: (a) Original Suicides, 30.06 (16); (b) Replication Suicides, 28.20 (15); (c) Potential Suicides, 28.75 (24); (d) Matched Controls, 19.06 (16); (e) Original Random Controls, 23.09 (116); (f) Replication Random Controls, 16.77 (30). It can be seen that the mean scores for the two groups of suicide Ss (and the unit selected potential suicides) are all quite distinct from those obtained for the three nonsuicidal groups. For all nine possible comparisons between suicidal and control groups (two suicide groups and one high potential group, each contrasted with the three nonsuicide samples), the differences obtained were significant beyond the $p < .001$ level. The results clearly support the validity of the revised SPS.

REFERENCE

- MISKIMINS, R. W., DECOOK, R., WILSON, L. T., & MALEY, R. F. Prediction of suicide in a psychiatric hospital. *Journal of Clinical Psychology*, 1967, 23, 296-301.

¹ An extended report of this study may be obtained without charge from R. W. Miskimins, Fort Logan Mental Health Center, 3520 West Oxford Avenue, Denver, Colorado 80236, or for a fee from the American Science Information Service. Order NAPS Document NAPS-00271 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001; remitting \$1.00 for microfiche or \$3.00 for photocopies. Requests for reprints should be sent to R. W. Miskimins at the above address.



NONPROFESSIONAL AND PROFESSIONAL HELPERS

two significant re-evaluations
forthcoming
from HRW

PSYCHOTHERAPEUTIC AGENTS

New Roles for Nonprofessionals, Parents,
and Teachers

Edited with commentaries by Bernard G. Guereney,
Jr., Rutgers, The State University

Through articles and editorial commentary, Professor Guereney explores remedial methods and procedures based on the strategy of using nonprofessionals as trained and supervised agents of the professional psychotherapist. The book considers in detail problems and solutions in selecting and training nonprofessionals, and includes detailed descriptions of therapeutic projects involving nonprofessionals.

April 1969

608 pages

\$9.95

HELPING AND HUMAN RELATIONS

A Primer for Lay and Professional Helpers
Volume I: Selections and Training

By Robert Carkhuff, University of Massachusetts,
Amherst

The first of a two-volume series, this volume deals creatively and pragmatically with the effective selection and training of lay and professional helpers for social work, counseling, clinical work, and psychiatry, as well as other areas of human relations.

August 1969

288 pages

\$7.50 (tent.)



Holt, Rinehart and Winston, Inc.

383 Madison Avenue, New York, New York 10017

*A new revision
of major significance . . .*

APA PUBLICATION MANUAL

1967 Revision

The APA PUBLICATION MANUAL has undergone a complete revision for the first time in 10 years. Although the revised edition prescribes the same basic format for manuscripts as the previous edition, major changes and additions were made in the text in an effort to clarify and simplify the instructions on how manuscripts should be prepared.

The APA PUBLICATION MANUAL is intended primarily as a guide for authors and editors in the preparation of material to be published in APA journals. However, in preparing the original edition it was recognized that the MANUAL could, and should, be written to accommodate a wider range of manuscripts. The revised PUBLICATION MANUAL has retained that original concept.

\$1.50

Order from:

AMERICAN PSYCHOLOGICAL ASSOCIATION
1200 Seventeenth Street, N. W.
Washington, D. C. 20036

EUROPE

To utilize the low group fares of the major airlines — especially in connection with the International Congress of Psychology in London July 27, and the Annual American Psychological Association Convention in Washington August 31, 1969 — we have scheduled several groups to London, the Continent, the Iberian Peninsula, Scandinavia, Yugoslavia and Greece. Round trip jet fares start at \$230 (New York/London.) Write to Division "P" Travel Wholesalers International, ITT Building, Washington, D.C. 20036, or call collect (202) 296-9161.

Consider this important
new text . . .

CASEBOOK ON ETHICAL STANDARDS FOR PSYCHOLOGISTS

The application of the 19 principles comprising the Code of Ethics for all psychologists is presented. Each principle is illustrated by actual, but disguised, cases involving alleged violations of the principle and action taken by the Committee on Scientific and Professional Ethics and Conduct.

Four Appendixes state and amplify the additional principles applicable to specific areas of psychology:

- Ethical Practices in Industrial Psychology*
- Automated Test Scoring and Interpretation Practices*
- Standards for APA Directory Listings of Private Practice*
- Guiding Principles for the Humane Care and Use of Animals*

The arrangement of the casebook into a general section of cases and four additional specific sections increases its versatility as a reference book for individuals and as a text for instructional purposes.

Price: \$1.00

Order From:

AMERICAN PSYCHOLOGICAL
ASSOCIATION
1200 Seventeenth Street, N. W.
Washington, D. C. 20036

The Difference

between
continuing
delinquency
and a good
citizen
is personal
interest and
individual care
under
professional
direction.

The Problem

America's most arrested
person is only 16
years of age.
His criminal
career
began at age 9.

It Can Be Dealt With

The St. Francis
Boys' Homes have
rehabilitated juvenile
delinquents since 1945.
From every state, of
every race and religion,
some as young as 12, others
as old as 18 . . . more than
80% have never been arrested
after graduation from SFBH.
The chain of delinquency
broken, they are citizens,
taxpayers, good husbands,
good fathers.

The Right Goal

Mere custody is not enough. The goal must be a change inside the young man himself — enough change to keep him out of trouble, enough change to enable him to live successfully in our society.

For Information write:

St. Francis Boys' homes

(Episcopal)

Admissions Secretary

(New England and Mid Atlantic States)

P. O. Box 630, Lake Placid, N.Y. 12946

(All Other States)

P. O. Box 1348, Salina, Kansas 67401

Journals of the AMERICAN PSYCHOLOGICAL ASSOCIATION

Subscription Rates for Calendar Year 1969

AMERICAN PSYCHOLOGIST. Official papers of the Association and articles on Psychology. Monthly. The 1969 volume is Vol. 24. Subscription: \$10.00 (foreign \$10.50). Single copy \$1.00.

CONTEMPORARY PSYCHOLOGY. Critical reviews of books, films, and research material in the field of psychology. Monthly. The 1969 volume is Vol. 14. Subscription: \$10.00 (foreign \$10.50). Single copy \$1.00.

JOURNAL OF ABNORMAL PSYCHOLOGY. Theoretical and research conditions in behavior pathology. Bimonthly; first issue appears in February. The 1969 volume is Vol. 74. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00. (Continues volume sequence of the former Journal of Abnormal and Social Psychology.)

JOURNAL OF APPLIED PSYCHOLOGY. Applications of psychology in business and industry. Bimonthly; first issue appears in February. The 1969 volume is Vol. 53. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF COMPARATIVE AND PHYSIOLOGICAL PSYCHOLOGY. Original contribution in the field of comparative and physiological psychology. Beginning in 1969: Monthly. Three volumes per year, four issues in each volume. The 1969 volumes are Vols. 67, 68 and 69. Subscription: \$30.00 for 3 vols. (foreign \$30.50). Single copy \$3.00.

JOURNAL OF COUNSELING PSYCHOLOGY. A primary publication medium for research in counseling theory and practice. Bimonthly; first issue appears in January. The 1969 volume is Vol. 16. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF CONSULTING AND CLINICAL PSYCHOLOGY. Research in clinical psychology: psychological diagnoses, psychotherapy, personality, psychopathology. Bimonthly; first issue appears in February. The 1969 volume is Vol. 33. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF EDUCATIONAL PSYCHOLOGY. Studies of learning and teaching; measurement of psychological development, psychology of school subjects, methods of instruction, school adjustment. Bimonthly; first issue appears in February. The 1969 volume is Vol. 60. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY. Original contributions of an experimental character. Monthly. Beginning 1969: Four volumes per year, three issues in each volume. The 1969 volumes are Vols. 79, 80, 81 and 82. Subscription: \$40.00 for 4 vols. (foreign \$40.50). Single copy \$4.00.

JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY. Theoretical and research papers on personality dynamics, group process, and the psychological aspects of social structure. Monthly. Three volumes per year, four issues in each volume. The 1969 volumes are Vols. 11, 12 and 13. Subscription: \$30.00 for 3 vols. (foreign \$30.50). Single copy \$3.00.

PSYCHOLOGICAL ABSTRACTS. Noncritical abstracts of the world's literature in psychology and related subjects. Monthly. Approximately 20,000 abstracts published annually. The 1969 volume is Vol. 43. Subscription: \$40.00 (foreign \$40.50). Single copy \$4.00.

PSYCHOLOGICAL BULLETIN. Evaluative reviews of research literature, discussions of research methodology in psychology. Monthly. Two volumes per year, six issues in each volume. The 1969 volumes are Vols. 71 and 72. Subscription: \$20.00 for 2 vols. (foreign \$20.50). Single copy \$2.00.

PSYCHOLOGICAL REVIEW. Original contributions of a theoretical nature. Bimonthly; first issue appears in January. The 1969 is Vol. 76. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

NEW IN 1969 DEVELOPMENTAL PSYCHOLOGY. Studies on growth and development and their associated variables. Bimonthly; first issue appears in January. The 1969 volume is Vol. 1. Subscription \$10.00 (foreign \$10.50). Single copy \$2.00.

Discounts: AGENTS and BOOK DEALERS receive a 10% discount on all orders. There is no club arrangement. There are no quantity, cash, library, or other special discounts. *Payment in U. S. funds.*

★ Subscriptions are available on a January through December basis only ★

AMERICAN PSYCHOLOGICAL ASSOCIATION
Subscription Department
1200 Seventeenth Street, N.W.
Washington D. C. 20036

JOURNAL OF CONSULTING and CLINICAL PSYCHOLOGY

June 1969

Vol. 33, No. 3

Contents

ARTICLES

- Behavior Therapy: Observations and Reflections: *Marjorie H. Klein, Allen T. Dittmann, Morris B. Parloff, and Merton M. Gill*..... 259
- Abstraction on Proverbs in Process-Reactive Schizophrenia: *Robert Cancro*..... 267
- Conservatism as a Predictor of Humor Preferences: *Glenn D. Wilson and John R. Patterson*..... 271
- College Students' Memories of Their Parents: A Factor Analysis of the CRPBI: *Herbert J. Cross*..... 275
- Differential Therapeutic Responding of A and B Quasi-Therapists to Schizoid and Neurotic Communications: *Juris I. Berzins and Edward Seidman*..... 279
- Experimental Approach to Defining the Role of Social Desirability in Personality Assessment: Is There One Response Process or Two?: *Stephen Leff and Nancy Lamb*..... 287
- Attitudes toward Death: A Psychological Perspective: *Herman Feifel*..... 292
- Sensing Humor: Latency and Amplitude of Response Related to MMPI Profiles: *Douglas S. Holmes*..... 296
- Psychiatric Hospital Experience and Attitudes toward "Mental Illness": *Jean J. Smith*..... 302
- Development of an Empathy Scale: *Robert Hogan*..... 307
- Personality Characteristics of Therapists: Description of Relevant Variables and Examination of Conscious Preferences: *Aron Z. Spilken, Martin A. Jacobs, James J. Muller, and Jane Knitzer*..... 317
- Aversion Techniques in Behavior Therapy: Some Theoretical and Metatheoretical Considerations: *G. Terence Wilson and Gerald C. Davison*..... 327
- Encoding of Attitude by a Seated Communicator via Posture and Position Cues: *Albert Mehrabian and John T. Friar*..... 330
- Motivational Patterns in the Families of Adjusted and Maladjusted Boys: *Edward J. Murray, Arthur Seagull, and David Geisinger*..... 337
- Reality Attentiveness-Inattentiveness and Externalization-Internalization in Defensive Style: *Marshall Bush, Robert Hatcher, and Martin Mayman*..... 343
- Validation Study of Marital Status and the Self-Report Scale for Process-Reactive Schizophrenia: *Donald H. Meichenbaum*..... 351
- Patterns of Performance on WISC Similarities in Emotionally Disturbed and Brain-Damaged Children: *Leon P. Hall and LaVerne LaDriere*..... 357
- An Attempt to Test Assumptions about Some Indications of Negativism on Psychological Tests: *Elizabeth Fox and Sidney J. Blatt*..... 365
- Sensation Seeking Scale as a Predictor of Need for Stimulation during Sensory Restriction: *John Hocking and Malcolm Robertson*..... 367
- Follow-Up of the Later Careers and Lives of 1,000 Boys Who Dropped Out of High School: *Starke R. Hathaway, Phyllis C. Reynolds, and Elio D. Monachesi*..... 370

BRIEF REPORTS

- Driver Accidents and the Neuropsychiatric Patient: *Matthew W. Buttiglieri, M. I. Charles E. Woodson, Marie Guenette, and Mae Thomson*..... 381
- Perceived Child-Rearing Practices and Self-Disclosure Patterns: *Joseph A. Doster and Bonnie Ruth Strickland*..... 382

LIST OF MANUSCRIPTS ACCEPTED..... 350, 356, 364

ERRATA..... 270

Brief Reports

The *Journal of Consulting and Clinical Psychology* will accept Brief Reports of research studies in clinical psychology for early publication without expense to the author. The procedure is intended to permit the publication of soundly designed studies of specialized interest or limited importance which cannot now be accepted because of lack of space. Several pages in each issue will be devoted to Brief Reports, published in the order of their receipt without respect to the dates of receipt of the regular articles. Most Brief Reports appear in the first or second issue to go to press following their final acceptance.

An author who wishes to submit a Brief Report:

1. Sends the Brief Report, limited to one printed page and prepared according to the specifications given below.
2. Also sends to the Editor a full report of the research study, in sufficient detail to give a clear account of its background, procedure, results, and conclusions, which will be filed with the National Auxiliary Publications Service (NAPS) to insure indefinite availability.
3. Prepares at least 100 mimeographed copies of the full report, which the author will send without charge to all who request it as long as the supply lasts.
4. Agrees not to submit the full report to another journal of general circulation.

Specifications

Brief Report. The Brief Report should give a clear, condensed summary of the procedure of the study and as full an account of the results as space permits.

To insure that the Brief Report will be no longer than one printed page, its typescript, in-

cluding all matter except the title and the author's lines, must not exceed 85 lines averaging 42 characters and spaces in length. Set the type-writer margins for short lines of 42 characters, which are 3.5 inches long in elite typing, and 4.2 inches long in pica.

The manuscript of the Brief Report must be *double spaced* throughout. Except for its short lines, it follows the standard style of the 1967 revision of the *APA Publication Manual*. Headings, tables, and references are avoided or, if essential, must be counted in the 85 lines. Each Brief Report must be accompanied by a footnote in the style below, which is typed on a separate sheet and *not* counted in the 85-line quota:

¹ An extended report of this study may be obtained without charge from John Doe (giving the author's full name and address) or for a fee from the American Society for Information Science. Order NAPS Document — from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001; remitting \$— for microfiche or \$— for photocopies.

Extended report. Because the extended report is intended for photoduplication, and is not copy to be sent to a printer, its style should differ in several ways from that of other manuscripts: (a) The extended report should be typed with single spacing for economy in duplication. (b) Tables and figures should be placed adjacent to the text which refers to them. A caption should be typed below each figure. (c) Footnotes should be typed at the bottom of the page on which reference is made to them. In other respects, the full report is prepared in the style specified by the *Publication Manual*.

Style of Manuscripts. Manuscripts must be prepared in the style described in the *Publication Manual of the American Psychological Association (1967 Revision)*, obtainable for \$1.50 from the Washington office of the APA.

Early Publication. Authors may secure the early publication of an acceptable article in the first issue to go to press after its acceptance by arranging for the article to be printed as extra pages, increasing the number of pages that subscribers receive. Information about charges for early publication may be obtained from the Managing Editor.

National Auxiliary Publications Service. To save printing costs, materials supplementary to articles may be deposited with the NAPS. To obtain copies, order by Document Number, as cited in the reference footnote, from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001. The indicated fee must be remitted in advance.

Abstracts. Manuscripts must be accompanied by an abstract of 100–120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or they may be obtained from the Editor or from the APA Central Office.

Copyrighted material. The following policies govern reprinting or other forms of reproduction of APA-copyrighted journal materials. Permission by the American Psychological Association must be requested for reprinting all tables, figures, and text of more than 500 words (total from one article) in length. There will be a charge of \$5.00 per page for such permission except when the requester is the author of the article to be reprinted, or the article is to be reproduced in a limited number of copies solely for instructional purposes.

Reprints. Authors receive 50 reprints gratis. Additional quantities may be ordered when returning proofs.

BEHAVIOR THERAPY:

OBSERVATIONS AND REFLECTIONS

MARJORIE H. KLEIN

Wisconsin Psychiatric Institute, University of Wisconsin

ALLEN T. DITTMANN, MORRIS B. PARLOFF

National Institute of Mental Health

AND MERTON M. GILL

Downstate Medical Center, State University of New York

This paper presents impressions based on 5 days' observation of the clinical activities of Joseph Wolpe and Arnold Lazarus at the Eastern Pennsylvania Psychiatric Institute. Behavior therapists now treat a wider range of patients and use a greater variety of treatment methods than before. Treatment is becoming longer and more complicated, with concomitant lowering of success rates. There are points where behavior therapy is not so simple and straightforward as the popular stereotype of it: clinical inference is essential in diagnosis, hierarchy selection, patient management, and ongoing evaluation. Elements of suggestion in the patient-therapist relationship may play a powerful role in treatment. Implications of these impressions for research and clinical training are discussed.

Like many clinicians and researchers today we have been impressed by claims made for behavior therapy: its superior effectiveness, straightforwardness, efficiency, flexibility, and objectivity. In order to learn more about how behavior therapists function we undertook to spend a period of time observing ongoing behavioral treatment under the guidance of two of its most prominent proponents, Joseph Wolpe and Arnold Lazarus. In this paper we wish to present some questions, comments, and inferences based on observations of their clinical activities which we were privileged to make during a typical week at the Eastern Pennsylvania Psychiatric Institute. Over 5 working days we saw both private and demonstration sessions for cases in various stages of treatment, and occasionally were able to interview patients. We had access to tape recordings of other treatment sessions, attended staff meetings and case conferences, and discussed many issues with both Wolpe and Lazarus. We are aware that 1 working week is in many ways an unusual length of time for outsiders to visit a clinical operation, but realize that a different week may have brought different patients and yielded other impressions.

At the outset we wish to express our appreciation for the extraordinary cooperation of our hosts and our deep respect for their unusual openness. Wolpe and his associates were both patient and helpful in their efforts to answer our questions, many of which we must admit were less than gentle and friendly and reflected our own psychoanalytically oriented training, experience, and biases. In the interests of accuracy in writing this paper, we sent drafts to both Wolpe and Lazarus, and they made a number of comments about it. Many of these concerned misstatements of fact on our part, and we have incorporated these by making the necessary corrections. Others were comments about interpretation, and we have, with their kind permission, included them in the article, separated from the body of our text by the use of smaller type size identifications of the commenters. We are greatly indebted to Wolpe and Lazarus for making this method of presentation possible.

Therapy as practiced by Wolpe and his colleagues is among the most widely used form of behaviorally oriented treatment. It was first described fully by Wolpe in 1958, with a more technique-oriented exposition including new developments recently available (Wolpe &

Lazarus, 1966). Complete unanimity of approach to theory and technique is not to be found in any so-called school of psychotherapy, and behavior therapists are no exception. Even the two we visited are quite different on some issues, Lazarus being a good deal more eclectic in his choice of techniques. Still, there are some basic general principles. The chief of these has to do with reciprocal inhibition, in which the fear response (or some other maladaptive response) to those situations which form the basis for the patient's complaint is inhibited by an incompatible response, usually a positive one. The positive responses most frequently used are relaxation and self-assertion. Relaxation is brought about by a modification of Jacobson's technique, and for each symptom the patient is instructed to imagine situations like those he fears, arranged in hierarchical order from least to most fear-producing. When the patient is able to imagine each situation and remain relaxed the therapist goes on to the next hierarchy scene. Where self-assertion is used as the new response, the patient is instructed in the relevant procedures for performance between sessions. Treatment is considered successful when all major symptomatic complaints are reorganized in this fashion.

Although most of what we encountered has been spelled out explicitly in the writings of Wolpe and others, some of what we saw surprised us, and led us to understand the literature more fully when we returned to it later. In addition, there were a few impressions which have not, as far as we know, been spelled out completely in any publication. We shall note these as we come to them and state them rather tentatively. We shall begin with a description of the range of patients who came in during our week of observation and of the techniques we saw employed and shall go on to discuss the implications of what we have seen for research in psychotherapy and for programs of clinical training.

Patients

Consistent with Wolpe's emphasis that his method has wide applicability, cases we observed under treatment went well beyond simple phobias or social anxieties (cases for whom behavior therapy has been thought to be

a specific) to include a broad range of complex psychoneurotic problems, character neuroses or borderline psychotic problems (acute psychoses are excluded). Many had been resistant to prior therapeutic efforts. To the clinician of ordinary persuasion a number of these patients would seem to have extremely poor prognosis and present a formidable challenge. They were not particularly gratifying patients for the behavior therapist either. These patients were often unwilling or unable to meet what are considered the basic requirements for behavior therapy or to perform correctly the therapeutic tasks that were assigned. In some of the more complex cases we observed (and in others we were told about) the patients were not optimally relaxed or would at times exaggerate or misreport their levels of relaxation. On occasions inquiry would reveal that the hierarchy scene actually "imagined" was considerably different from that which the therapist had presented. As a consequence new cues or response dimensions were often introduced, reducing or heightening anxiety levels spuriously. In addition, compliance with the very specific instructions for response performance outside the therapy setting seemed equally difficult at times to obtain or to verify.

In general, it was our impression that broadening the clientele for behavior therapy to include "less good" cases has introduced inevitable strain on the method and potential depression of its outcome rate. It should be noted that behavior therapists have not by plan increased the range of problems they treat. Rather, as they have become better known, they have been referred a broader spectrum of patients, a number of whom, as we have said, had experienced unsatisfactory results in the care of others.

Techniques

From our acquaintance with the literature we knew intellectually at least that behavior therapists do not work in a unitary fashion and indeed take pains to vary their approach from case to case. We were surprised to find, however, that within most cases, too, a number of manipulations were routinely employed. Even where desensitization was the primary technique, others such as assertiveness or ex-

pressiveness training, manipulation of behavior outside the treatment setting, and education in learning principles were also included. With the more complex cases the spectrum of techniques became even broader. In addition to desensitization involving a variety of hierarchies, many variants of assertion and expression, often elaborate role-playing and behavioral programs of homework were devised to correct response deficiencies. Along with these specific procedures we found also that patients were given a good deal of indoctrination, teaching, and exhortation, apparently intended to provide a rationale for the treatment and to enhance motivation.

These remarks may well sound like complaints by outsiders that behavior therapists use a "shotgun" approach. But such an accusation would be unfounded, since the many techniques do have a unity in the behavior therapist's view in that they are thought either to produce reciprocal inhibition of anxiety or to alter stimulus-response contingencies. It would make no more sense to insist that the behavior therapist be limited to one behavioral manipulation than it would to insist that another kind of therapist restrict himself to one content area or interview technique. It is appropriate, however, to note the apparent contradiction between the proliferation of methods in behavior therapy and the popular conception, based partly on hope, partly on the behavior therapists' writings, that this is a simple and straightforward treatment of the neuroses. It is also appropriate for us to look into the research implications of a many-method approach, as we shall do later on in this paper.

The popular notion of simplicity in behavior therapy covers not only treatment techniques but diagnostic procedures as well. Many people suppose that the therapist begins by clearly and systematically defining the patient's problems in terms of manageable hierarchies and then selects appropriate responses to be strengthened or weakened. We found little support for this conception of behavior therapy diagnosis in our observations. Indeed the selection of problems to be worked on often seemed quite arbitrary and inferential. We were frankly surprised to find the presenting symptomatic complaint was often

sidestepped for what the therapist intuitively considered to be more basic issues. Most surprising to us, the basis for this selection seemed often to be what others would call dynamic considerations. The distinction between "secondary" (that is, the superficial) and "primary" (the more basic, underlying) was even made openly on occasion by behavior therapists: The words we have put in quotes or their terms. The literature, of course, gives no hint of this development.

JW: This conveys the impression that the writers assume that "more basic" in a behavioral sense is the same thing as the psychoanalysts mean. This is really quite incorrect. For example, it may be found that a person's fear of enclosed spaces is based on a fear of developing certain physical symptoms that may lead, as he believes, to his death, and he avoids enclosed spaces because once within them he cannot easily run out and seek help. It is the fear of his symptoms that would then be the primary focus of this person's treatment. What is involved here is the pinpointing of the time stimuli to anxiety, but these are still extrinsic, and no reference is made to "intrapsychic mechanisms."

AAL: One must be careful when using a term like "dynamic trends." A behavior therapist may trace a patient's anxiety to faulty identification with an inadequate father coupled with excessive sibling rivalry, compounded by ambivalence engendered at the hands of an inconsistent mother. However, these basic and formative interpersonal encounters are then taken at face value and do not presuppose additional "dynamics" such as "castration anxiety," "incestuous wishes" or "death wishes," etc. As to the statement, "The behavior therapy literature, of course, gives no hint of this development," may I refer you to Lazarus (1966).

We also learned that behavior therapists prefer to structure hierarchies on dimensions of time and space (for example, "a month before an examination," and leading up to the time of an examination; "standing in a hall a mile away from the hospital" and gradually approaching it; treading water in a pool 5 feet deep and 5 feet from the edge," working up to greater depths and distances). The examples of hierarchies reported in the behavior therapy literature are generally much more flexible and lifelike than the ones we encountered during our visit.

One point which the behavior therapists' writing makes quite clear but which did not strike us until our visit has to do with the ongoing evaluation of the patient progress.

Because behavioral treatment is posited as highly specific, it follows that success depends on the patient's exact and close cooperation with the therapist's instructions. It is therefore very important for the therapist to test this cooperation repeatedly during treatment. The therapist must also constantly assess the patient's progress on hierarchy dimensions. With the possible exception of role-playing, the therapist is dependent in this evaluation upon the patient's report both of progress outside treatment and of events within the sessions themselves. Since there are no independent procedures for evaluating or verifying his report, the patient has considerable leeway to bias his report in order to please, frustrate, or otherwise manipulate the therapist, or to meet some personal expectation. And the form of much of the feedback from the patient (i.e., lifting his finger if he feels an increase in anxiety during desensitization or doing nothing if he does not) gives the therapist very few cues for distinguishing valid from invalid reports. Thus the therapist must use considerable intuition to assess progress and correct the treatment plan. This all serves to highlight a very basic discrepancy between the theoretical orientation of behavior therapy and its actual practice. While the theory clearly calls for the manipulation of overt behavior, the therapist typically deals primarily with the patient's report of his image of that behavior.

JW: We all have "images" of everything we talk about. If a patient says that he has a feeling of anxiety, this is presumed to be his image of some responses within him. Except insofar as this kind of image cannot be shared by another observer, there is no difference between such a report and the report of something perceived in the world outside the subject. Either *could* be false: but what would motivate patients to mislead their helpers? And what insinuit would make them all lie according to the same rules with regard to their emotional responses to repetitions of an image—that is, they report either that anxiety is weak and progressively diminishes, or that it is strong and does not?

AAL: It is incorrect to state that, "the theory calls for manipulation of overt behavior." All forms of behavior are covered by the theory, including "thinking behavior."

Perhaps the most striking impression we came away with was of how much use behavior therapists make of suggestion and of how

much the patient's expectations and attitudes are manipulated. Behavior therapists are not at all silent on this point in their descriptions of technique, but the literature did not prepare us for the unabashed suggestions that therapists directed toward their patients. The major arena for suggestion is in the orientation period of treatment. Here the therapist tells the patient at length about the power of the treatment method, pointing out that it has been successful with comparable patients and all but promising similar results for him too. The patient is provided with a detailed learning-theory formulation of the etiology of his problems and is given a straightforward rationale for the way in which the specific treatment procedures will "remove" his symptoms. The patient's motives and values may also be considered so as to "correct misconceptions" which block desirable courses of action or restrict the effect of treatment. Indeed it seemed to us that treatment plans and goals were laid out in such a detail that the patient was taught precisely how things would proceed and what responses and changes were expected of him all along the way. A quite complete account of these procedures may be found in Wolpe and Lazarus (1966, pp. 16–20).

AAL: Both Wolpe and I have explicitly stated that relationship variables are often extremely important in behavior therapy. Factors such as warmth, empathy, and authenticity are considered necessary but often insufficient.

Although Wolpe and other behavior therapists are reluctant to ascribe therapeutic effectiveness to these features of the relationship it is difficult for us to believe that they do not constitute an important part of the treatment. Certainly the explicit, positive, and authoritative manner with which the therapist approaches the patients seems destined if not designed to establish the therapist as a powerful figure and turn the patient's hopes for success into concrete expectations. The introductory education in learning theory, in addition, must function to make the treatment more plausible and provide a simple and coherent frame of reference for the patient's understanding of his difficulty. Further, the focus of the treatment philosophy on the role

of external psychonoxious environmental factors in the formation of the patient's problems must be quite reassuring for many patients. At the very least the therapist's evident willingness to assume major responsibility for correcting the patient's problems may be especially important in helping him overcome inertia and start on the path to change.

JW: Paul's (1966) study in which psychoanalytically oriented therapists obtained significantly better results with desensitization than with either their own insight methods or with another clearly structured method indicates a relatively small role for the factors suggested here.

The Course of Treatment

As behavior therapists have gained more experience with their methods and behavior therapy has been more widely used, a number of complementary trends seem to have emerged; techniques have become more polished, tending to shorten the treatment of specific cases, but at the same time patients with more complicated problems have been seeking behavior therapy, tending to lengthen treatment. In the first years of Wolpe's experience with behavior therapy the average number of interviews per case rose from 25 (Wolpe, 1952, p. 827) and 26 (Wolpe, 1954, p. 217) to 45 (Wolpe, 1958, p. 218), but declined sharply thereafter: Lazarus reports an average of 14 interviews per case in 1963 (Lazarus, 1963, p. 75). At the present time it is our impression that the length of treatment is on the increase again. Some of the patients we observed had been in treatment for over a year. In others the number of interviews per week had increased dramatically. One patient was being seen four times a week, and this had been going on for approximately 9 months. We felt that these trends reflected the shift in the case load to include more complex or seriously disturbed cases. Also related are changes over the years in reported outcome rate. From the over 90% improvement originally reported (see Wolpe, 1952, 1954, 1958) the claims have been modified to approximately 80% (Lazarus, 1963), with some of Wolpe's associates now reporting 70% rate. These figures cannot be taken literally as they were not derived from data specifically designed to yield comparative outcome rates—indeed the

last figure is purely anecdotal. What we wish to convey by referring to them is that behavior therapy seems to be taking longer and showing less spectacular successes. Why this should be so could be the topic of extensive study. It should be noted, however, that these trends are familiar ones among psychotherapies as they are introduced and find their way.

One final impression about the course of treatment: It is interesting to note that the ongoing argument between behaviorist and other methods concerning symptom substitution seems to be taking a new turn. When confronted with a patient who is successfully treated for his initial array of problems, but continues to report disabling complaints, the therapist talks not of symptom substitution but of faulty or inadequate treatment. It may be argued, for example, that (a) the original formulation of the case did not consider the correct stimulus contingencies; (b) "secondary" anxieties may have masked the more "primary" or situational pervasive anxieties; or (c) the patient did not comply with the instructions or work hard enough to extend his learning outside the treatment setting. While it may be commendable for behavior therapists to assume primary responsibility for slow or uneven progress in their patients, the distinction between backsliding due to inadequate treatment and symptom substitution is certainly not as obvious to outsiders (especially to dynamically oriented ones) as it is to behavior therapists. In any case it is a distinction which is difficult to verify.

JW: Symptom substitution means replacement of one symptom by another. It is not symptom substitution if a patient complains of symptoms that he had before treatment began, though other symptoms have been removed by therapy. "Backsliding" is in no sense commonly observed.

AAL: "Symptom substitution" is an unfortunate term. It presupposes that unresolved unconscious complexes result in the terminal manifestation of new problems. Patients can always acquire new aberrant responses (e.g., when faced with inimical life situations). Is this "symptom substitution?"

Research Implications

In this section we shall consider a number of implications our observations have for be-

havior therapy research. It should be clear, and let us stress it if it is not, that these problems are by no means unique to research in behavior therapy. Indeed, what struck us most forcefully was the fact that the issues and research problems which confront behavior therapy are the same familiar ones which have long troubled all psychotherapy research.

The first and most obvious research problem behavior therapists must face is a consequence of the diversity of their techniques. While it is true, as we said earlier, that behavior therapists believe all of their techniques are united in their derivation from theories of learning, the question of the exact function and effectiveness of each technique is still an empirical one. Research questioning which of the many techniques is best suited to which patient, and which produces which result, would teach behavior therapists how to make their treatment more effective and efficient.

In making these points let us stress again that we are not criticizing behavior therapy for its technical variety; the range of patients seen certainly calls for flexibility in treatment, and the underlying experimental and empirical philosophy demands constant technical adjustments. Nonetheless, it remains that the question of the theoretical utility of a many-faceted method, whatever its theoretical persuasion and whatever its clinical value, has never been adequately dealt with in psychotherapy research. Thus the fact of variety in behavior therapy makes it as difficult to talk of behavior therapy as a single method of treatment to be evaluated and compared with others as it is to talk of a single method of insight therapy. The subtleties and variations within methods may be as important as the unifying theories. It may be possible to compare therapeutic approaches on some dimensions—behavioral symptoms versus unconscious conflicts as the “core” of neurosis, for example, in comparing behavior and insight therapies—but it must be remembered that the dimensions are theoretical ones. They may relate only indirectly to the specific therapist activities during the treatment session, and there may be equally important variables that may be overlooked because they do not happen to be in the theoretical spotlight.

In our earlier section on technique we emphasized the role of suggestion in behavior therapy practice. We did this because we were surprised at the extent and consistency of its use. The research implications are obvious: Even though these factors are not called techniques, in that they do not follow from the theoretical framework, they nevertheless have an effect on the patient, and these effects must somehow be understood and differentiated from the effects of the theoretically important “techniques.”

JW: Suggestion is nothing but the use of words to obtain the responses that are associated with these words. If deliberate evocation of responses in this way brings about habit change, it is just as much behavior therapy as habit change through responses otherwise contrived. Words are probably used with this intent more consciously by behavior therapists than most other therapists.

AAL: The statement that suggestion as a technique does not follow from the theoretical framework is not correct. “Suggestion” can only have an effect by virtue of responses already in the person's repertoire (i.e., previously conditioned cues) which the suggestions come to elicit. Furthermore, if suggestion enables the person to attempt new responses, these may have positive effects. One thus endeavors quite deliberately to maximize the “placebo effect.”

Again, this problem is not unique to behavior therapy, but again, it is the popular notion that behavior therapy is simple and straightforward, behaviorally oriented, and thus untrammelled by these side effects. In practice, all therapists make suggestions, by telling their patients about the methods they employ, and by explaining why they believe these methods will be helpful. The particular behavior therapists we observed were experienced clinicians (whose training included dynamic methods), and they went about their business very confidently and skillfully. Much of what they did was what any clinician does in dealing with patients, and they did it by second nature, so to speak. The point is that all of these activities, the second-nature ones, the overt suggestions, the specific “technique”—all must be examined for their effects on the patient.

AAL: Recent experimental results lead me to endorse the authors' contentions here. Indeed, even the results of a specific technique like systematic desensitization cannot be accounted for solely in terms of graded hierarchies and muscle relaxation.

The more complex cases highlight some other difficult questions for assessing the outcome of behavior therapy. It is generally assumed that the behavior therapist, basing his judgment on change in manifest symptoms, enjoys a simpler criterion of outcome than does the "dynamic" therapist whose criteria are cast in such inferential terms as conflict resolution or structural personality change. But the apparent objectivity of symptomatic change is lost when the behavior therapist treats patients who present complex symptom pictures. Unless all symptoms are successfully removed, one must decide in assessing outcome whether all symptoms are to be equally weighted, whether only those symptoms mentioned at intake should be counted, or whether complaints emerging later should be given secondary status. Questions like these are certainly not new to behavior therapists, any more than they are to therapists of other persuasions. In fact, it looks as if they may have already begun to work toward some answers although perhaps unwittingly: The differentiation between "primary" and "secondary" problems which we mentioned earlier may be a first try at ordering symptoms for therapeutic attack. That this differentiation was made on the dynamic basis was our inference from our observations. Whatever the basis, it must be defined more clearly before it can be used in designing research studies.

One more research implication of the complex cases: Behavior therapists have relied almost exclusively on patient reports, both for evaluating the success of manipulations within a given session and for assessing the outcome of the entire case. In general it was our impression that the more disturbed patients were especially unlikely to give reliable reports. The problem seemed to make it difficult for therapists to handle these patients clinically. It brings about insuperable difficulties in comparing patients with each other in a research study, or for that matter in comparing two reports collected at different times from the same patient. The obvious way out of this problem is to rely more heavily on external measures, preferably behavioral ones, as criteria for the success of treatment.

Implications for Clinical Practice and Training

At present there is a clearly discernible trend in a number of training institutions to emphasize behavior therapy considerably, both in theory and technique. We are concerned that such training practices not provoke an overreaction against "dynamic" theory and practice. It may be well to recall that the innovators of behavior therapy have each had considerable training and experience with classical psychodynamic orientations.

JW: It seems to be suggested here that previously acquired psychoanalytic knowledge plays an integral part in the practical repertoire of behavior therapists. Earlier passages have implied that because the behavior therapist is not always content with "a simple mechanical acceptance of the patient's statements regarding his fears and anxieties," and may go into probing activities for the sake of clarification to establish the antecedents, this is *ipso facto* psychoanalytic. It is in a sense *analytic*, but in every important respect repugnant to whatever is specific to the psychoanalytic theory of neurosis, and different from the practices of psychoanalytic therapy.

Despite their present enthusiasm for behavior therapy the fact remains that previously acquired skills still form an integral part of their repertoire. The effect is perhaps most clearly evident in the selection of patient-relevant hierarchies, which we have stressed before. Recall also our observations that Wolpe and his associates make very effective use of the patient-therapist relationship to establish a context in which the specific behavioral techniques can be utilized most effectively. Although the behavior therapist would not wish to ascribe his results to any "dynamic" theories he may have learned in the past, it is clear that many clinical decisions in the treatment are based on an understanding of some functional organization of behavior, and an appreciation of the power of the relationship.

The danger for training programs as we see it is that the trainee will be exposed principally to learning theory and behavioral techniques without being adequately trained in other areas of psychology. It is especially those clinical skills of interviewing and diagnosis, and those theoretical skills that grow with knowledge of a wide range of therapeutic

theories that prevent training from becoming sterile and one-sided. Without these skills the training will severely limit the competence of the therapist and restrict the quality of research and theoretical development possible within the field. Training programs which do not specifically attempt to develop a broad range of clinical skills are in danger of reducing the effectiveness of the entire profession.

JW: With regard to the statement that behavioral training should avoid being one-sided, the answer is that a very good knowledge of all kinds of reactions and interactions is highly desirable. Our position is that it is not important to know a great deal about theories, but enough to obtain a historical perspective for present-day practices. From the point of view of the problems presented by patients, only propositions that have some scientifically meaningful, factual support are worth anything at all. If behavior therapists have tended to ignore propositions from the psychoanalytic field, it has been because of a dearth of them that qualify for respect on this criterion.

AAL: Do we need only a knowledge of general psychology and therapeutic theories to be effective clinicians or do we need in addition a course in sensitivity training and rational thinking combined with a mélange of specific techniques?

Little benefit to the field of psychotherapy is to be gained from developing yet another closed system of treatment. We believe that the stimulus provided by the behavior therapist's apparent willingness to examine the con-

sequences of his procedure may be a positive one for the field. Perhaps ultimately the greatest contribution the behavior therapist can make to psychotherapy theory and practice is his dedication to making both the process and the outcome of treatment as objective and as efficient as possible. If the discrepancies between behavior therapy and other therapies can be subjected to empirical tests, then our ultimate aim of achieving a more effective theory of personality change may also be served.

REFERENCES

- LAZARUS, A. A. The results of behaviour therapy in 126 cases of severe neurosis. *Behaviour Research and Therapy*, 1963, 1, 69-79.
- LAZARUS, A. A. Broad-spectrum behaviour therapy and the treatment of agoraphobia. *Behaviour Research and Therapy*, 1966, 4, 95-97.
- PAUL, G. L. *Insight vs. desensitization in psychotherapy*. Stanford, California: Stanford University Press, 1966.
- WOLFE, J. Objective psychotherapy of the neuroses. *South African Medical Journal*, 1952, 26, 825-829.
- WOLFE, J. Reciprocal inhibition as the main basis of psychotherapeutic effects. *A.M.A. Archives of Neurology and Psychiatry*, 1954, 72, 205-226.
- WOLFE, J. *Psychotherapy by reciprocal inhibition*. Stanford, California: Stanford University Press, 1958.
- WOLFE, J., & LAZARUS, A. A. *Behavior therapy techniques: A guide to the treatment of neuroses*. Oxford: Pergamon Press, 1966.

(Received May 28, 1968)

ABSTRACTION ON PROVERBS IN PROCESS- REACTIVE SCHIZOPHRENIA¹

ROBERT CANCRO²

The Menninger Foundation

Fifty-one consecutive, drug-free, male admissions, with a senior-staff diagnosis of schizophrenia, were given the Phillips Prognostic Rating Scale (PRS) and the Benjamin Proverb Test shortly after entering the hospital. There was no relationship between classification by PRS and abstraction when drug effects and duration of current hospital stay were controlled. Using only those cases who had no previous history of psychiatric hospitalization, thereby controlling for previous and current hospitalization, still failed to produce a significant relationship. The suggestion of a racial effect emerged. This is discussed along with certain theoretical implications of the finding.

The work of Goldstein (Bolles & Goldstein, 1938; Goldstein, 1944; Goldstein & Scheerer, 1941) suggested the hypothesis that schizophrenics showed a particular impairment of the abstract attitude, which could be differentiated from the characteristic impairment of an organic. More recently there have been a number of studies which have reported that process schizophrenics show a greater impairment of the abstract attitude than reactives (Becker, 1956; Herron, 1962; Johnson, 1966; Judson & Katahn, 1963; Meadow, Greenblatt, Funkenstein, & Solomon, 1953; Meadow, Greenblatt, & Solomon, 1953).

Johnson (1966) comments on a number of methodological deficiencies in the earlier studies, highlighting particularly the importance of using a reliable means of classification and controlling for the effect of vocabulary differences. There are even more serious methodological problems which have to be considered. The use of chronically institutionalized and drugged patients introduces an indeterminate amount of variability. Reported differences in abstract attitudes between process and reactive cases can be a function of group differences in type, dosage, and responsiveness to drugs. The effects of long-term institutional care are such that reactive cases who usually have been in the hospital for a

shorter period of time cannot be compared to process cases who have been in the hospital for longer periods of time. Chronic institutionalization is very similar to a subclinical sensory deprivation experience, and it would not come as a surprise if abstracting ability on the proverbs were to be affected by duration of hospital stay. The present study was undertaken to compare abstracting ability on proverbs in process and reactive schizophrenics on whom these variables were controlled.

METHOD

Subjects

Fifty-one consecutive male admissions to the psychiatric observation wards of a large city hospital, who had a senior-staff diagnosis of schizophrenia, ranged in age from 18 to 44 and were fluent in English, were admitted to this study. The age range was selected in order to screen out early adolescent reactions, involuntional states, and premature senility. Patients whose clinical picture was complicated by organic, neurological, or extraneous factors, such as alcoholism, were rejected. Patients were not given tranquilizers or long-acting sedatives throughout the period of testing and examination so that performance on the test of abstraction would not be contaminated by drug effect.

Procedure

Toward the end of the first week of hospitalization, the patient was given the 11 proverbs from the Benjamin Proverb Test (Benjamin, 1944). He was informed that the results of the study would not play a role in determining his length of hospital stay. The proverbs were orally administered with instructions for the patient to give the most general answer that he could. If he responded with a concrete or only partially abstract answer it was accepted, and

¹ Supported in part by the Commonwealth Fund of New York, through a Research Fellowship held by the author at the State University of New York, Downstate Medical Center.

² Requests for reprints should be sent to Robert Cancro, The Menninger Foundation, Topeka, Kansas 66601.

he was encouraged to try to give a more general meaning. This was done in order to maximize the patient's abstract response and to compensate for individual or group differences in initial motivational level. The proverbs were scored on the Meadow Scale (Meadow, Greenblatt, Funkenstein, & Solomon, 1953) by an independent judge who had no knowledge of the patient's process-reactive classification. The total score for the 11 proverbs was used as the measure of abstraction.

Patients were interviewed, and the material obtained was used to fill out the Prognostic Rating Scale (PRS; Phillips, 1953), which served to rate the patients on the process-reactive continuum. Patients were not divided into nominal categories on the basis of PRS, but rather were assigned their actual score for the calculation of Pearson product-moment correlations. The use of consecutive schizophrenic admissions raised concerns about the uncontrolled effects of variables such as age, education, race, Wechsler Adult Intelligence (WAIS) vocabulary, and history of previous psychiatric hospitalization. Correlation coefficients between these variables and the PRS score were determined. As an additional precaution the population was divided so that patients with PRS scores in the range of 2-14 were grouped as reactive ($N = 20$), and those with a PRS score between 15 and 32 ($N = 31$) were grouped as process (Phillips, 1953). The two groups were then compared to see if they were adequately matched. Race was included as a variable because differences in the quality of education resulting from school segregation might affect the abstraction score.

RESULTS

There are no significant correlations between the PRS score and age ($r = -.13$), years of education ($r = .16$), race ($r = .16$), WAIS scaled vocabulary score ($r = .19$), and history of previous psychiatric hospitalization ($r = .10$). The matching of the process and reactive groups on the cited demographic variables is presented in Table 1. There are no significant group differences. Of the total sample, 54.9% of the patients had no history of prior psychiatric hospitalization. The slightly higher incidence of previous hospitalization in the reactive group is not significant ($\chi^2 = 1.071$, $df = 1$), yet it raises some concern about the representativeness of the sample. The total group was followed for a period of 3 years after admission and the PRS did successfully predict future length of hospital stay for both process and reactive cases (Cancro & Sugerman, 1968). The mean duration of previous hospitalization in the reactive group was 4.25 months as compared

TABLE 1
DEMOGRAPHIC MATCHING OF PROCESS AND REACTIVE GROUPS

Item	Process	Reactive
Mean age	28.3	31.6
Mean years of education	11.3	11.1
Mean scaled vocabulary (WAIS) score	9.4	9.1
Proportion of nonwhite patients	29.0%	30.0%
Previous mental hospitalization	40.0%	60.0%
First mental hospitalization	60.0%	40.0%
Mean duration in months of previous hospitalization	9.06	4.25

Note.—For the process group, $N = 31$; for the reactive group, $N = 20$.

to the process group mean of 9.06 months. While this difference is statistically significant ($t = 2.86$, $p < .005$, one-tailed), it is not a very large one.

The correlation between the actual score on the PRS and the abstraction score is .005 for the total sample. The mean abstraction score for the process group is 6.0 and for the reactive group is 7.2; these scores do not differ significantly. The use of recent admissions controlled for length of current hospital stay but did not control for duration of previous hospitalization. Taking those 28 patients who were admitted to a psychiatric facility for the first time and dividing them on the basis of their PRS scores into process and reactive categories revealed no mean differences in abstraction. The mean abstraction score for the reactive group was 6.3, while the mean for the process group was 6.1. The correlation between the actual PRS score and the abstraction score on the 28 first admissions was not significant.

There is a relationship in this sample between race and history of prior psychiatric hospitalization. The white patients are more likely to have been hospitalized before ($\phi = .35$, $p < .02$, two-tailed). There are weak relationships between race and age ($r = .26$, $p < .10$, two-tailed), and between race and abstraction ($r = .25$, $p < .10$, two-tailed). The white patients tend, although not significantly, to be older and to have higher scores on abstraction. Restricting the sample to nonwhite patients ($N = 15$) brings out a weak relationship between scores on the

PRS and abstraction in the expected direction ($r = -.35$, $p < .10$, one-tailed). The group means between nonwhite process and reactive cases differ weakly in the expected direction ($t = 1.63$, $p < .10$, one-tailed).

DISCUSSION

The failure to demonstrate a difference in abstract ability between process and reactive cases, when drugs and duration of current hospital stay are controlled, raises many questions. Virtually all state hospital patients diagnosed as schizophrenics are on tranquilizers. Those patients whose clinical pictures are more malignant, that is, the process cases, are most likely to receive large doses of the major tranquilizers. If having a slender body build is related to poor prognosis in schizophrenia, then we would expect process cases, as a group, to be getting even higher effective levels of drugs in terms of milligrams of tranquilizer per kilogram of body weight per day. Reported linear relationships between abstract ability and normal, reactive, and process groups may be a function of differences in group levels of drug ingestion.

While the current hospital stay was controlled, there was a difference between groups in mean duration of total previous hospitalization. It would be difficult to attribute the failure in demonstrating a difference in level of abstraction to the effect of a mean of 5 months' additional lifetime hospitalization. There is even less reason to assume that the slightly longer prior hospitalization of the process group would raise their abstract score. Nevertheless, the correlation between duration of previous hospitalization and abstraction was obtained and it was not significant.

The results suggest the usefulness of controlling for possible racial effects on proverb tests of abstraction in schizophrenics. The quality of education, as well as its quantity, must be controlled if we are to use tests which are sensitive to educational differences. The greater disorganization of the process schizophrenic may manifest itself in the proverbs when the poorer quality of education prevents the "reflex" correct response, thereby producing apparent group differences in abstract ability.

The value of proverbs tests in general as

measures of the abstract attitude in schizophrenia is open to substantial doubt. The abstraction score is determined in part by the correctness of the response and therefore indirectly by the absence of intrusions of schizophrenic thought. The proverbs were re-scored by a different independent judge, who had no knowledge of the abstraction or Phillips scores, for frequency of signs of schizophrenic thought disorder, using a scale developed by Cancro (1962). The correlation between abstraction and frequency of signs of thought disorder was $-.38$, $p < .005$. The greater the number of intrusions, the lower the abstract score. This finding supports the suggestion that proverb measures of abstraction are weighted by the relative "cleanness" of the response. It is strongly recommended that new methods of measuring the abstract attitude that are not so vulnerable to educational and psychopathological differences be developed.

REFERENCES

- BECKER, W. C. A genetic approach to the interpretation and evaluation of the process-reactive distinction in schizophrenia. *Journal of Abnormal and Social Psychology*, 1956, 53, 229-236.
- BENJAMIN, J. D. A method for distinguishing and evaluating formal thinking disorders in schizophrenia. In J. S. Kassin (Ed.), *Language and thought in schizophrenia*. Los Angeles: University of California Press, 1944.
- BOLLES, M., & GOLDSTEIN, K. A study of the impairment of "abstract behavior" in schizophrenic patients. *Psychiatric Quarterly*, 1938, 12, 42-65.
- CANCRO, R. A comparison of process-reactive schizophrenia. Unpublished doctoral dissertation, State University of New York, Brooklyn, 1962.
- CANCRO, R., & SUGERMAN, A. A. Classification and outcome in process-reactive schizophrenia. *Comprehensive Psychiatry*, 1968, 9, 227-232.
- GOLDSTEIN, K. Methodological approach to the study of schizophrenic thought disorder. In J. S. Kassin (Ed.), *Language and thought in schizophrenia*. Los Angeles: University of California Press, 1944.
- GOLDSTEIN, K., & SCHEERER, M. Abstract and concrete behavior. *Psychological Monographs*, 1941, 53(2, Whole No. 239).
- HERRON, W. G. Abstract ability in the process-reactive classification of schizophrenia. *Journal of General Psychology*, 1962, 67, 147-154.
- JOHNSON, M. H. Verbal abstracting ability and schizophrenia. *Journal of Consulting Psychology*, 1966, 30, 275-277.
- JUDSON, A. J., & KATAHN, M. The relationship of autonomic responsiveness to process-reactive

schizophrenia and abstract thinking. *Psychiatric Quarterly*, 1963, 37, 19-24.

MEADOW, A., GREENBLATT, M., FUNKENSTEIN, D. H., & SOLOMON, H. C. The relationship between the capacity for abstraction in schizophrenia and the physiological response to autonomic drugs. *Journal of Nervous and Mental Disease*, 1953, 118, 332-338.

MEADOW, A., GREENBLATT, M., & SOLOMON, H. C. "Looseness of association" and impairment in abstraction in schizophrenia. *Journal of Nervous and Mental Disease*, 1953, 118, 27-35.

PHILLIPS, L. Case history data and prognosis in schizophrenia. *Journal of Nervous and Mental Disease*, 1953, 117, 515-525.

(Received for early publication October 23, 1968)

ERRATA

In the article "Size Constancy in Retarded Versus Normal Children: A Developmental Hypothesis" by Bruce J. Richman, Harold N. Kellner, and Dean A. Allen (*Journal of Consulting and Clinical Psychology*, 1968, 32, No. 5, pp. 579-582), a statement appearing below Table 1 is erroneous. Reanalysis has indicated that none of the between-groups comparisons of the means presented in that table are significant at $p < .05$. The results reported, therefore, must be regarded as inconclusive and subject to caution in interpretation.

In the article "Hypnotist and Manner of Presentation Effects on a Standardized Hypnotic Susceptibility Test" by D. Eugene Thorne and Ernst G. Beier (*Journal of Consulting and Clinical Psychology*, 1968, 32, No. 5, pp. 610-612), there is a misspelling of J. V. Cabibi's name on page 611 and again in the reference list. The reference should read: Cabibi, J. V., Hughes, H. H., & Butler, J. R. Behavioral effects on the hypnotic process by a change in hypnotist. *Journal of Clinical Psychology*, 1966, 21, 334.

In the article "Relationship of Personality Characteristics to Persistence in Psychotherapy" by Ralph E. Van Atta (*Journal of Consulting and Clinical Psychology*, 1968, 32, No. 6, pp. 731-733), on page 731, column 2, line 23, delete the phrase "and by Sullivan, Miller, and Smelzer (1958) . . ." After line 26 add the following: "Sullivan, Miller, and Smelzer (1958) confirmed the association of educational and occupational level with both length of stay and improvement in therapy."

CONSERVATISM AS A PREDICTOR OF HUMOR PREFERENCES

GLENN D. WILSON¹

AND

JOHN R. PATTERSON

Institute of Psychiatry, London

University of Canterbury, New Zealand

Of a sample of 139 high school students, humor preference profiles were compared for 30 Ss scoring highest on the Conservatism Scale, and 30 Ss scoring lowest. As predicted, high-conservative Ss tended to prefer "safe," formal types of humor (e.g., puns), while low conservatives (liberals) expressed greater appreciation of transparently "libidinal" types of humor (sick and sexual). Results were interpreted as indicating the importance of individual differences in the extent to which it is necessary for appetitive content to be disguised by the structural aspects of the joke medium before humorous affect is evoked.

Freud (1905) emphasized the distinction between tendentious humor, which has the function of giving expression to sexual and aggressive impulses (broadly defined), and innocent or nontendentious humor, which derives its effect solely from the techniques and conventions of the joke medium itself. In psychoanalytic theory, tendentious humor is successful only so long as the "joke facade" is effective in disguising the underlying libidinal intent from the consciousness of the audience, otherwise it evokes discomfort or disgust instead of humorous pleasure.

Experimental studies (Gollob & Levine, 1967; Singer, Gollob, & Levine, 1967) appear to support the psychoanalytic notion that an important function of humor is that of providing a socially acceptable setting for the partial expression of normally prohibited impulses, and that intensification of the inhibitions relating to these impulses may interfere with the appreciation of humor. These have so far dealt only with situationally induced inhibitions relating to aggression, although the investigators do contemplate the possibility of "personality variables which determine differential susceptibility to this type of interference [Singer et al., 1967, p. 568]."

The Conservatism Scale (C Scale; Wilson & Patterson, 1968) seems promising in this respect since it may be regarded as measuring the extent to which "parental" attitudes and prohibitions have been internalized. High scorers on the C Scale may be assumed, on construct grounds at least, to be intolerant,

resistant to change, authoritarian, dogmatic, and antihedonistic in outlook. The present experiment tests the hypothesis that persons scoring high on conservatism, having chronically mobilized inhibitions, will be less appreciative than persons scoring low on conservatism of humor involving the transparent expression of sex and aggression, but not of humor which is either purely formal or in which libidinal content is comparatively veiled. In other words, we are postulating a difference between liberals and conservatives in the proportion of the nontendentious component (amount of facade) which is required before a joke is capable of evoking humorous affect.

METHOD

The six categories of humor selected for study are listed, with definition and examples, in Table 1. It is not claimed that this classification of humor is in any sense comprehensive, merely that it provides a good sample of humor types differing along the "formal-libidinal" continuum.

As the next step in the construction of a humor preference test, a search was made of magazines for cartoons which appeared to correspond fairly unambiguously to the six categories. The category definitions were then given to four graduate students who were asked to sort the cartoon pool in terms of them, and those cartoons which were reliably sorted into the same category were retained. Finally, five cartoons corresponding to each category were selected to be included in the test booklet.

The Ss were 139 high school students (75 female and 64 male), aged 15-19. High school students are well suited to a study of this kind because they comprise a fairly homogeneous group with respect to age, marital status, education, and occupation. The cartoon booklet was given to Ss with instructions to rate each cartoon according to the scale: (1) Not at all funny, (2) Slightly funny, (3)

¹ Requests for reprints should be sent to Glenn D. Wilson, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London, S.E.5, England.

TABLE 1
CATEGORIES OF HUMOR EMPLOYED

Category	Definition	Example
Puns	Play on the sound and meaning of words	A carrier staggering under the weight of a grandfather clock is asked: "Have you got the time on you?"
Incongruity	Familiar elements appearing out of usual context	A butcher with an animal-plan erected on a music stand is applying his hacksaw to a side of meat in the manner of a double-bass player.
Antiradical	Ridicule of radical or Bohemian elements in society	As they pass by a hardware store, a beatnik is shielding his girl-friend's eyes from the sight of a bath and washbasin.
Antiauthority	Ridicule of conventionally respected persons and institutions	A truck-driver waiting at a stop-light is casually stubbing out his cigarette on the helmet of a motorcycle cop also waiting beneath him.
Sick	Delight in the morbid, gruesome, and sadistic	A little girl is swinging joyfully on the legs of a man who has just hanged himself.
Sexual	Participation in, or anticipation of, sexual pleasure	A playboy attired in silk dressing gown is welcoming an attractive blonde. "Come in, Miss Faversham. Looks like you're the only one to come to the party."

Moderately funny, (4) Very funny, (5) Extremely funny. Personality information was gathered by subsequent administration of the C Scale, and the EPI (Eysenck & Eysenck, 1964).

RESULTS AND DISCUSSION

Puns and incongruity jokes are fairly exclusively formal, cognitive, and respectable in

content, while the sick and sexual categories constitute the outright expression and indulgence of aggressive and sexual impulses, respectively. The antiradical and antiauthority categories can both be described as forms of social satire involving the ridicule of certain classes of persons and institutions, a type of indirect, interpersonal aggression.

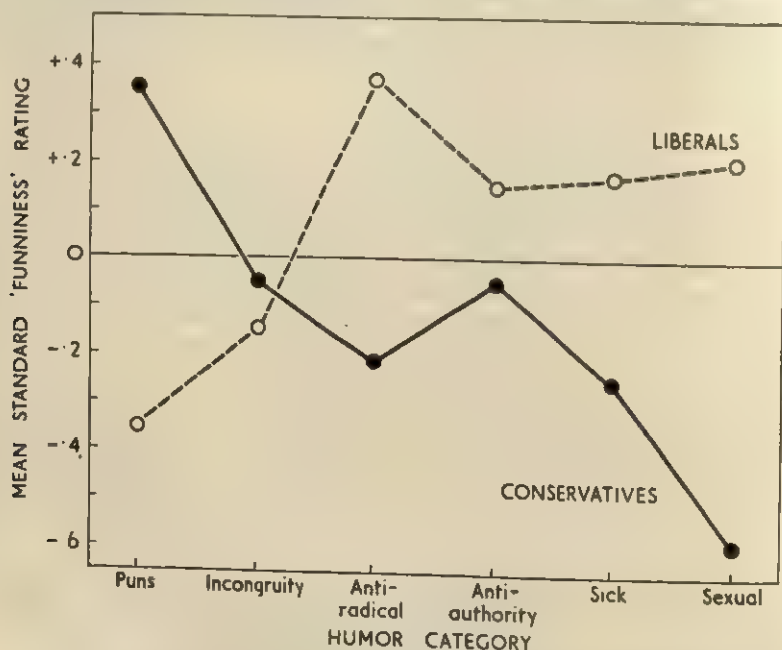


FIG. 1. Comparison of humor preference profiles for the 30 most conservative and 30 most liberal Ss. (The six categories of humor are ordered along a dimension of "tendentiousness" on an a priori construct basis.)

TABLE 2
PRODUCT-MOMENT CORRELATIONS BETWEEN CONSERVATISM SCORES AND "FUNNINESS"
RATINGS FOR EACH OF THE SIX HUMOR CATEGORIES

Ss	Puns	Incongruity	Antiradical	Antiauthority	Sick	Sexual
Males	.35**	.14	-.29*	-.08	-.10	-.28*
Females	.26*	.08	-.20	-.03	-.15	-.23*
Total group	.27**	.10	-.22**	-.10	-.18*	-.24**

Note.—For Males, $N = 64$; for Females, $N = 75$; Total $N = 139$.

* $p < .05$.

** $p < .01$.

Figure 1 shows a comparison of humor preference profiles for the 30 Ss scoring highest on the C scale (conservatives) and the 30 Ss scoring lowest (liberals). The general pattern of the profiles is very much as predicted, with conservatives tending to prefer the very "safe" or nontendentious types of humor and liberals expressing a greater liking for the more openly "libidinal" or tendentious types of humor. Using a t test, these differences are significant beyond the .01 level except for the incongruity and antiauthority categories, which are not significant. Correlations between conservatism scores and humor ratings on each of the six cartoon categories are given for males and females separately, and for the total group, in Table 2.

It might be argued that a self-report technique for measuring humor preferences is susceptible to the influence of social desirability, and that such a factor may have mediated the relationship found between conservatism and "funniness" ratings. The evidence which is available, however, does not support this argument. The EPI "lie scale" is essentially a measure of social desirability or "faking good," and these scores bear little relationship to the humor scores, the only significant correlation being that between the lie score and the pun category ($r = .18$, $p < .05$). In any case, there is suggestive evidence that persons scoring high on the C Scale tend to give more socially desirable responses ($r = .20$, $p < .05$), consistent with their general tendency to "play it safe" (Patterson & Wilson, in press).

Of the other EPI personality variables, extraversion was significantly related only to sexual jokes ($r = .31$, $p < .01$), while neu-

roticism did not relate to any of the humor scales.

Although there were certain complex differences between males and females in the humor ratings, these were independent of the effects of conservatism since neither age nor sex was significantly related to conservatism.

The one deviation from the expected ordering of differences between liberals and conservatives, the partial reversal on the antiradical and antiauthority categories, is interesting. Satire directed at radical targets (e.g., beatniks, artists, protest demonstrations) was rated higher by liberals than conservatives, though it is the latter who are supposedly antagonistic toward such institutions. On the other hand, there was no significant difference between the two groups in the expressed enjoyment of humor based on the ridicule of authoritarian institutions (e.g., Church, law) which ostensibly gain respect from conservative elements. Psychoanalytically, this finding might be interpreted as suggesting that conservatives, although superficially respectful of authority, are actually harboring a certain amount of resentment at some "unconscious" level, which is expressed and revealed in their unexpectedly high appreciation of antiauthority humor. This expression of antiauthority feelings through humor preference is presumably only possible because the joke mechanism provides an effective distraction from the aggressive and "antisocial" content of the cartoons.

Taken generally, the results of this experiment suggest that there are considerable individual differences in the extent to which it is necessary for sexual and aggressive content of a joke to be disguised before humorous

affect is evoked. In fact, with very liberal persons there may be little or no necessity for any such veiling or mitigation of "impulse expression." The tendency for high-conservative Ss to rate the purely formal humor types (particularly puns) higher than liberals might then have arisen through a process of compensation for low ratings on the more threatening categories, based on a need to exhibit some "sense of humor."

REFERENCES

EYSENCK, H. J., & EYSENCK, S. B. G. *The Eysenck Personality Inventory*. London: University of London Press, 1964.

FREUD, S. *Jokes and their relation to the unconscious*. (Ed. & trans. by J. Strachey) New York: Norton, 1960 (Orig. publ. 1905).

GOLLOB, H. F., & LEVINE, J. Distraction as a factor in the enjoyment of aggressive humor. *Journal of Personality and Social Psychology*, 1967, 5, 368-372.

PATTERSON, J. R., & WILSON, G. D. Anonymity, occupation, and conservatism. *Journal of Social Psychology*, in press.

SINGER, D. L., GOLLOB, H. F., & LEVINE, J. Inhibitions and the enjoyment of aggressive humor. *Journal of Personality*, 1967, 35, 562-569.

WILSON, G. D., & PATTERSON, J. R. A new measure of conservatism. *British Journal of Social and Clinical Psychology*, 1968, 7, 264-269.

(Received May 6, 1968)

COLLEGE STUDENTS' MEMORIES OF THEIR PARENTS:

A FACTOR ANALYSIS OF THE CRPBI¹

HERBERT J. CROSS²

University of Connecticut

The revised 18-scale form of the Children's Reports of Parental Behavior Inventory was administered to 119 college females and 99 males. Four factor analyses—females' reports on fathers, females' reports on mothers, males' reports on fathers, and males' reports on mothers—were performed. Three factors, corresponding to Schaefer's three dimensions of Acceptance vs. Rejection, Psychological Autonomy vs. Psychological Control, and Firm Control vs. Lax Control, emerged. Factor structures were similar in all analyses except males' reports of mothers. These differences are discussed in terms of the relative importance of dimensions used by students to evaluate their parents' behavior.

Schaefer (1965a, 1965b) has done extensive work on conceptualizing and measuring children's reports of parental behavior. He has modified his two-dimensional circumplex model of maternal behavior (Schaefer, 1959) to now (Schaefer, 1965b) include three dimensions: Acceptance versus Rejection, Psychological Autonomy versus Psychological Control, and Firm Control versus Lax Control. This model of parent behavior is roughly similar to those independently developed by Roe (1957), Slater (1962), and Becker (1964) and seems to summarize a great deal of thinking and data about parents' behavior toward their children.

Much of the data on parental behavior has been gathered from children's reports (Schaefer, 1965b; Slater, 1962). The parent-child relationships and the influence of the parent can in this way be studied through the offspring's description, which enormously simplifies data collection. Several studies (e.g., Lyle & Levitt, 1955; Serot & Teevan, 1961) have pointed up the usefulness of the offspring-perception technique and one (Schaefer & Bayley, 1967) has directly af-

firmed the validity of certain Children's Reports of Parental Behavior Inventory (CRPBI) scales.

The present study is an attempt to extend Schaefer's findings (1965b) to another norm group and to explore the usefulness of the revised 18-scale form of the CRPBI.

In order to support Schaefer's model, these data should factor in a way that can be reasonably explained by the three dimensions of Acceptance-Rejection, Psychological Autonomy-Psychological Control, and Firm Control-Lax Control.

METHOD

Subjects

The Ss were 119 females and 99 males, mostly freshmen and sophomores who were volunteers (for credit) from general psychology classes at the University of Connecticut.

CRPBI

This form of the CRPBI consists of 18 scales, 6 of which are based on 16 items and 12 of which are based on 8 items. Items consist of behaviors, for example, "Isn't very patient with me," "Says I'm a big problem," "Tells me I'm good looking," which S indicates are like, somewhat like, or not like the parent being described. These responses are scored 3, 2, and 1, respectively, and summed to yield scores on individual scales.

Procedure

The Ss were tested in classroom-size groups. Separate, but virtually identical, mother and father forms were administered in a single session, in counter-balanced order. The Ss were asked to describe their parents' behavior as they remembered it when they were 16 years of age.

¹ This research was supported by the University of Connecticut Research Foundation.

² Requests for reprints should be sent to Herbert J. Cross, Department of Psychology, University of Connecticut, Storrs, Connecticut 06268. The cooperation of the University of Connecticut Computer Center is gratefully acknowledged. Roy E. Shore deserves thanks for his helpful comments on the manuscript. The assistance of Earl Schaefer and Herbert Kaufman with data analysis and interpretation is gratefully acknowledged.

Analysis

Four 18×18 (scales) product-moment correlation matrices were formed; female Ss' scores on the mother form, females on the father form, males on the mother form, and males on the father form. Each matrix was separately factor analyzed by the principal components method, after which an orthogonal Varimax rotation was performed, using communality estimates from the squared multiple-correlation coefficients in the diagonals. The program (Dixon, 1965) allows one option and that is the value which limits the number of eigenvalues of correlation coefficients used in the analysis. With the optional value set at 1.00, three factors resulted in three of the analyses; in the fourth analysis a fourth factor was extracted. However, its eigenvalue was only 1.006, it accounted for only 6% of the variance, and it appeared to be a specific factor, having a substantial loading only on the "Inconsistent Discipline" scale. Therefore, a three-factor solution was computed. The first three factors are essentially the same in both cases.

RESULTS

Table 1 presents the rotated factor loadings for both factor analyses done on females' reports of their mothers and fathers, respectively.

With regard to the analysis of reports of

mothers' behavior, the three factors extract 74% of the total matrix variance. Factor I, which accounts for 50% of the common variance, loads high on Acceptance, Child Centeredness, Positive Involvement, and Acceptance of Individuation. This factor shows high negative loadings on Rejection, Hostile Detachment, Hostile Control, Instilling Persistent Anxiety, and Withdrawal of Relations. Factor I is clearly a love versus hostility factor and it is nearly identical to Schaefer's Acceptance versus Rejection dimension (1965b, p. 554). Factor II accounts for only 26% of the common variance and shows high positive loadings on Control and Enforcement with high negative loadings on Non-Enforcement, Lax Discipline, and Extreme Autonomy, and is quite similar to Schaefer's third dimension, Firm Control versus Lax Control. Factor III accounts for 23% of the common variance and shows highest loadings on Possessiveness, Intrusiveness, Control through Guilt, Hostile Control, and Instilling Persistent Anxiety, and is similar to Schaefer's second dimension, Psychological Autonomy versus Psychological Control.

TABLE 1
ROTATED MATRICES FOR FEMALES' REPORTS OF
MATERNAL AND PATERNAL BEHAVIOR

Scale	Mother factors				Father factors			
	I	II	III	h^2	I	II	III	h^2
Acceptance	93	-16	00	90	89	-03	-28	87
Child Centeredness	82	-18	29	79	89	-06	-06	81
Possessiveness	30	-07	82	76	64	04	53	69
Rejection	-87	19	25	85	-71	13	59	87
Control	-21	69	45	73	-18	79	37	80
Enforcement	-45	70	27	76	-33	72	36	76
Positive Involvement	85	-11	15	76	85	01	06	72
Intrusiveness	07	34	70	61	29	36	61	58
Control through Guilt	-51	06	70	74	-16	04	85	74
Hostile Control	-69	32	55	88	-43	38	71	83
Inconsistent Discipline	-41	-42	40	50	-24	-54	48	58
Non-Enforcement	04	-85	00	73	-08	-89	-05	81
Acceptance of Individuation	88	-27	-06	85	79	-26	-36	83
Lax Discipline	18	-87	15	82	25	-79	04	70
Instilling Persistent Anxiety	-64	21	57	79	-41	28	71	75
Hostile Detachment	-88	07	09	79	-78	-02	49	85
Withdrawal of Relations	-63	03	40	56	-39	02	64	56
Extreme Autonomy	19	-63	-23	50	-07	-65	-12	44
Percentage of common variance	50	26	23		40	28	31	

Note.—Decimals omitted.

TABLE 2
ROTATED MATRICES FOR MALES' REPORTS OF
MATERNAL AND PATERNAL BEHAVIOR

Scale	Mother factors				Father factors			
	I	II	III	h^2	I	II	III	h^2
Acceptance	03	89	-02	80	90	-09	-07	83
Child Centeredness	29	84	05	79	86	12	14	77
Possessiveness	73	26	-01	60	28	05	78	69
Rejection	43	-68	-06	66	-72	03	43	70
Control	60	12	-60	73	-08	75	40	74
Enforcement	47	-08	-64	64	-19	76	23	68
Positive Involvement	15	85	01	75	88	00	12	79
Intrusiveness	70	09	-21	54	01	26	69	54
Control through Guilt	80	-22	05	70	-23	04	78	66
Hostile Control	83	-35	-19	84	-49	26	68	78
Inconsistent Discipline	27	-28	61	52	-44	-50	46	65
Non-Enforcement	-08	-02	88	78	-07	-85	03	73
Acceptance of Individuation	-25	79	20	72	87	-17	-15	82
Lax Discipline	07	16	85	75	15	-87	11	80
Instilling Persistent Anxiety	81	-12	-06	67	-43	20	74	77
Hostile Detachment	31	-81	10	76	-85	04	30	83
Withdrawal of Relations	49	-45	23	50	-60	06	45	58
Extreme Autonomy	-25	17	77	68	07	61	27	46
Percentage of common variance	36	37	28		43	27	29	

Note.—Decimals omitted.

With regard to the factor analysis of females' reports of father's behavior, each factor is remarkably similar to those just described. Each scale loads at approximately the same magnitude except for minor differences which seem to indicate that, while Factor III is substantially the same as in the first analysis, the factor axis is positioned at a slightly different location. A configurational analysis of the scales in three-dimensional space (cf. Schaefer, 1965b) shows essentially the same pattern as found in the first analysis.

Thus, an overall comparison between the two factor analyses indicates that they are quite similar in loading patterns and amounts of common variance accounted for by each factor. Further, in both analyses, the factor patterns are clearly analogous to Schaefer's proposed dimensions, with one minor difference. Schaefer's Psychological Autonomy versus Psychological Control factor is stronger than his Lax Control versus Firm Control factor while the findings here are reversed.

The factor structure of the males' reports of father data is strikingly similar to both the analyses presented in Table 1. The magni-

tude and direction of each loading is approximately the same as in the analyses of the data from females. Again the overall results clearly fit Schaefer's model.

The factor structure of males' reports of mothers' behavior data is clearly different from the other three analyses. The first factor accounts for slightly less common variance than the first factor in the other analyses (36% as compared to 43%, 40%, and 50%) and shows high loadings on Possessiveness, Intrusiveness, Control by Guilt, Hostile Control, and Instilling Persistent Anxiety. Factor I in this case is not an Acceptance versus Rejection factor but rather clearly emerges as a Psychological Autonomy versus Psychological Control factor. Factor II, which accounts for 26% of the common variance, factors out as the Acceptance versus Rejection factor and, in fact, the loading pattern is strikingly similar to that of Factor I in each of the other three analyses. Factor III, which accounts for 22% of the common variance, shows high *negative* loadings on Control and Enforcement, and high positive loadings on Inconsistent Discipline, Non-Enforcement, Lax Dis-

cipline, and Extreme Autonomy. Thus, it emerges as the Lax Control versus Firm Control factor and is similar to Factor II in the other three analyses.

DISCUSSION

Three of the four factor analyses presented here (females' reports of mothers and fathers, and males' reports of fathers) are highly similar. In each case Factor I, the Acceptance versus Rejection dimension, accounts for considerably more common variance than either Factor II or Factor III. It is, therefore, the most relevant dimension for females reporting on their mothers, females reporting on their fathers, and males reporting on their fathers. The dimensions of Psychological Autonomy-Psychological Control and Firm Control-Lax Control are clearly secondary to Acceptance-Rejection.

Mothers of males, in strong contrast, are evaluated primarily on the Psychological Autonomy versus Psychological Control Dimension, and secondarily on the Acceptance versus Rejection dimension. These two factors are nearly equal in amount of common variance they account for, however.

An obvious interpretation of the discrepancy between males' reports of their mothers and other reports is that the mother-son relationship is more psychologically intense, that is, mothers are more neurotically involved with sons than fathers are, or than fathers or mothers are with their daughters.

Even so, probably the most important overall dimension is Acceptance versus Rejection since it accounts for substantial amounts of variance in each analysis.

Some kind of dimensionalizing is probably useful for investigating the effects of parental behavior on personality, and Schaefer's three dimensions seem to be a step in the right direction.

This study leads to the following conclusions: (a) The 18-scale form of the CRPBI is a useful instrument for assessing young adults' memories of their parents' behavior. (b) These reports can be conceptualized effectively with Schaefer's three dimensions, Acceptance versus Rejection being the most relevant of the three. (c) Males' reports of their mothers are based more on the dimension of Psychological Autonomy versus Control than are males' reports of fathers, or females' reports of either mothers or fathers.

REFERENCES

- BECKER, W. C. Consequences of different kinds of parental discipline. In M. L. Hoffman & L. W. Hoffman (Eds.), *Review of child development research*. Vol. 1. New York: Russell Sage, 1964.
- DIXON, W. J. (Ed.) *BMD biomedical computer programs*. Los Angeles: Health Sciences Computing Facility, School of Medicine, UCLA, 1965.
- LYLE, W. H., JR., & LEVITT, E. E. Punitiveness, authoritarianism, and parental discipline of grade school children. *Journal of Abnormal and Social Psychology*, 1961, 63, 493-503.
- ROE, A. Early determinants of vocational choice. *Journal of Counseling Psychology*, 1957, 4, 211-217.
- SCHAEFER, E. S. A circumplex model for maternal behavior. *Journal of Abnormal and Social Psychology*, 1959, 59, 226-235.
- SCHAEFER, E. S. Children's reports of parental behavior: An inventory. *Child Development*, 1965, 36, 413-424. (a)
- SCHAEFER, E. S. A configurational analysis of children's reports of parent behavior. *Journal of Consulting Psychology*, 1965, 29, 552-557. (b)
- SCHAEFER, E. S., & BAYLEY, N. Validity and consistency of mother-infant observations, adolescent maternal interviews, and adult retrospective reports of maternal behavior. *Proceedings of the 75th Annual Convention of the American Psychological Association*, 1967, 2, 147-148.
- SEROT, N., & TEEVAN, R. C. Perception of the parent-child relationship and its relation to child adjustment. *Child Development*, 1961, 32, 373-378.
- SLATER, P. E. Parental behavior and the personality of the child. *Journal of Genetic Psychology*, 1962, 101, 53-68.

(Received May 9, 1968)

DIFFERENTIAL THERAPEUTIC RESPONDING OF A AND B QUASI-THERAPISTS TO SCHIZOID AND NEUROTIC COMMUNICATIONS¹

JURIS I. BERZINS² AND EDWARD SEIDMAN

University of Kentucky

This experiment investigated the proposition that the differential effectiveness of "A" and "B" therapists with schizoid and neurotic patients is based on behavioral differences which emerge even with untrained Ss in the appropriate interpersonal contexts. In a 3×2 factorial design, 72 male undergraduates (As, ABs or middles, and Bs) responded to tape-recorded schizoid and neurotic communications by writing self-chosen "helpful" responses. Analyses of these responses confirmed the predicted Therapist \times Patient interaction effects. When As were paired with the schizoid and Bs with the neurotic patient (relative to the opposite pairings), they gave longer responses ($p < .005$), emitted more responses of a declarative rather than questioning form ($p < .025$), and showed more positive ($p < .025$) and fewer negative ($p < .05$) Bales social-emotional reactions. The Ss intermediate on the A-B scale (ABs) did not react differentially to the two patient types.

Following clinical studies which demonstrated that so-called "A" therapists achieved greater success in psychotherapy with schizophrenic patients than "B" therapists (see the reviews by Betz, 1967; Carson, 1967) while Bs obtained better outcomes with neurotic patients than As (McNair, Callahan, & Lorr, 1962), several therapy-analogue studies have been reported which yielded results consonant with those of clinical studies, even though the "therapists" were totally untrained undergraduate males (Berzins & Seidman, 1968; Carson, Harden, & Shows, 1964; Sandler, 1965). The separation of therapists into A and B "types" has usually been accomplished by selecting extreme groups of responders to a small number of items taken from the Strong Vocational Interest Blank, the A-B scale, whose content ostensibly entails varying degrees of interest in manual-technical activities (Bs endorse these interests; As do not).

¹ This research was partially supported by a University of Kentucky summer research fellowship to the first author, and a National Institute of Mental Health predoctoral research fellowship (1-F1-MH-36960-01) to the second author. The authors gratefully acknowledge the use of the University of Kentucky Computer Center, and also express appreciation to John V. Haley for statistical consultation.

² Requests for reprints and copies of the A-B scale as well as of all stimulus materials should be sent to Juris I. Berzins, Department of Psychology, University of Kentucky, Lexington, Kentucky 40506.

While the relationship of such interests to performance in dyadic interactions of a "helpful" nature is only beginning to be understood, the finding that trained and untrained "therapists" perform similarly with the schizoid and neurotic classes of patients suggests that the A-B variable may have some "basic" bearing upon personality theory and measurement.

The present experiment was designed to investigate some features of the "natural" treatmentlike behaviors and reactions of college students varying in A-B status when they are exposed to communications representing two classes of patients, schizoid and neurotic. A portion of the present study, which has been reported previously (Berzins & Seidman, 1968), showed that, when As had responded "helpfully" (in writing) to a schizoid patient, and when Bs had responded to a neurotic patient, they reported greater satisfaction with their responses and found it easier to choose helpful responses than when As and Bs had responded under the opposite conditions of pairing. These differences in subjective reactions, which were in line with expectations based on clinical studies, left unanswered the question: On what *behavioral* differences are these interaction effects based?

This analogue experiment thus attempts to pinpoint, under controlled circumstances, dif-

ferences in the "helpful" behaviors of untrained As and Bs. At the same time, the freedom given to Ss (i.e., to emit responses they felt would be most helpful at each point that the taped communications were interrupted) makes the analytic task formidable. Whatever technique of analysis the investigator adopts, however, should yield results consistent with the prediction that instead of main effects for therapist or patient "type," Therapist \times Patient interaction effects should be obtained. In other words, it is expected that "better" responding takes place when As are paired with the schizoid patient, and Bs with the neurotic patient, than under the opposite Therapist \times Patient pairings.

METHOD

Design and Procedure

The essential features of this study have been reported previously (Berzins & Seidman, 1968). Briefly, the basic design was a $3 \times 2 \times 2$ factorial, using three independent groups of "quasi-therapists" (As, ABs or middles, and Bs), two patient types (AVOS, or avoidance of others, and TAS, or turning against the self) as representative of schizoid versus neurotic psychopathology (Phillips & Rabinovitch, 1958), and two orders of presentation (TAS-AVOS and AVOS-TAS). This report focuses upon the initial combinations of patient and "therapist"; second-position data are explored only in a limited fashion. Analyses were conducted on the resulting 3×2 factorial arrangement with cell frequencies randomly balanced at 12 (total $N = 72$).

The data were collected from undergraduate males attending their usual laboratory sections in introductory psychology at the University of Kentucky; Es were their usual laboratory instructors ($N = 8$). After listening to each patient communication segment, Ss were given 1 minute to give a "helpful" response (in writing) as though speaking directly to the patient. The Ss were classified into As, ABs (middles), and Bs on the basis of a trichotomization of the distribution of scores on a 15-item version of the A-B scale obtained in a laboratory meeting 1 week later.

Patient-Type Stimulus Materials

The tape-recorded, enacted patient communications and their validation have been presented in the previous report.³ Here, the first (of five) patient-communication segments are presented for illustrative purposes:

³ Again, the authors thank James E. Dublin and Robert D. Welch for their skillful enactment of the TAS and AVOS scripts.

TAS patient, Segment 1: The thing I keep thinking over and over . . . I should get going but I can't . . . it's just useless for me to try anything . . . I haven't been sleeping at all for a week . . . well, you know, I get some sleep but not enough to keep me going, not what I should . . . even with Sally lately I've been acting more like a jerk, a phony . . . and the more I think about it the more I get all tensed up . . . how long can I keep this up . . . I mean it just keeps going and on . . .

AVOS patient, Segment 1: Well I know this sounds a little crazy . . . and I know I'm not crazy but . . . it seems like things happen to me in a different way or . . . more than other people . . . or they just don't understand . . . I guess I mean wouldn't understand because I can't tell them what it is . . . like . . . the other day I was shaving and I looked in the mirror and the thought . . . sort of came to me like it was from the mirror, that the person in the mirror wasn't me at all . . . was somebody who shaves every day like everybody . . . and it was a funny feeling like I was him . . .

The TAS communications revolved around neurotic-depressive complaints, including suicidal ideation, while the AVOS tape, on the whole, showed a mixture of schizoid and paranoid features.

Dependent Measures and Analyses

It had been initially hoped that the self-chosen "helpful" responses of these Ss might be analyzable through the application of scales developed by various investigators for rating the content of therapy interactions (e.g., Rogers, Gendlin, Kiesler, & Truax, 1967; Strupp, 1957). It became apparent, however, that such scales were inappropriate for the analysis of the written responses of untrained Ss. Accordingly it was decided to follow a strategy which was content-free to varying degrees, in the hope that it would be the largely structural and formal characteristics of the "helpful" communications which would clarify the behavioral differences, if any, among the therapist-types. Data were analyzed according to a scheme which arranged the analyses in increasing order of reliance upon the content of responses and concomitantly required increasing degrees of inference from judges. The dependent measures were:

1. Length of response—a simple word count was obtained for each response.⁴
2. Form of response—responses were dichotomously coded as "declarative formulations" or "others" (largely questions). A declarative formulation was grammatically defined as any statement which did or could "naturally" end in a period or exclamation mark, and which in whole or in part had the patient as its direct referent (explicitly, e.g., using the

⁴ Special thanks are extended to Wynona J. Moore, Lois M. Gridley, and Peggy A. Mull for their painstaking transcription of the data on individual index cards.

second-person pronoun, or implicitly). This dichotomy of course did not discriminate between positively and negatively toned communications, but coding, performed without knowledge of Ss' A-B status, involved minimal inference and reliability was perfect.

3. Interaction characteristics of response⁵—the Bales Interaction Process Analysis (Bales, 1950) was employed according to standard instructions but with one modification—only one of the 12 Bales categories was assigned to an S's response during one response period. This procedure yielded five category scores per S. The Bales system is exhaustive and one-half of the categories permits classification on the basis of the affective characteristics of response. Scoring was performed by the writers, working independently and without knowledge of Ss' A-B status. While only one score was assigned per response, note was kept of additional (overlapping) categories. Initial agreement between scorers was 82%. Responses on which disagreement had occurred were resolved by consensus and one final score assigned to each response. Subsequently, the 12-category scores were collapsed into the four major classes of Positive (social-emotional) Reactions, Attempted Answers, Questions, and Negative (social-emotional) Reactions.

4. Content scales—three 5-point semantic-differential like scales were used. The scales were defined by three adjectives at the endpoints and comprised ratings of Relevance (relevant, accurate, to the point—irrelevant, inaccurate, beside the point), Acceptance (accepting, warm, friendly—rejecting, cold, unfriendly), and Assertiveness (assertive, active, strong—tentative, passive, weak). These ratings were independently performed by four advanced graduate students in clinical psychology⁶ who were given transcripts of the patient communications preceding Ss' responses. To control for possible biases associated with judges' own A-B status, two judges used were As and two were Bs. The reliabilities of the content ratings, calculated according to an analysis of variance technique (Winer, 1962, p. 124), were satisfactory for Relevance ($r_k = .70$) and Acceptance ($r_k = .79$), while agreement on Assertiveness was more modest ($r_k = .54$). The mean of the judges' ratings was used as the score for each S.

RESULTS

A-B Status and Length of Response

Three by two analyses of variance were performed on the number of words emitted in the "helpful" responses at each of the five points of interruption (hereafter termed trials), as well as on overall means. In the

⁵ The authors are indebted to Jesse D. Geller and Gordon F. N. Fearn for recommending the Bales analyses for these data.

⁶ The authors wish to thank Anthony Burry, Michael L. Campbell, Andrew Sappington III, and David J. Sheskin for serving as judges in this study.

latter analysis, no main effects for therapist or patient type were found, but a highly significant Therapist \times Patient interaction effect was observed ($F = 6.29$, $df = 2/66$, $p < .005$). The mean number of words per trial, presented in Table 1, reveal that "compatible" pairings (As with AVOS and Bs with TAS) elicit considerably longer responses than incompatible pairings (As with TAS and Bs with AVOS). For each patient, curves plotting mean words across trials showed no overlap between As and Bs, and the interaction F ratios were significant for four of the five trials. ABs (middles) were indeed intermediate in length of response and, more importantly, showed no differential responsiveness to the patient types.

It appears that, despite the simplicity of the dependent measure, the results attest to genuine behavioral differences in As and Bs as a function of the patient with whom they are paired. These differences were maximal on the very first trial, that is, after less than 1 minute of listening to the patient (interaction $F = 8.55$, $df = 2/66$, $p < .001$). The magnitude of the differences is noteworthy also: On the first trial with the AVOS patient, for example, As responded with twice as many words as did Bs (As: $M = 21.25$; Bs: $M = 10.00$).

A-B Status and Form of Response

For this analysis, all "helpful" responses were dichotomized into declarative formulations and "others." An analysis of variance for the number of declarative formulations emitted in five trials showed no main effects but, as in the preceding analysis, the Therapist \times Patient interaction effect was significant.

TABLE 1
MEAN NUMBER OF WORDS EMITTED
PER TRIAL

Therapist type	Patient type	
	TAS	AVOS
A	17.18	21.90
AB	20.02	20.10
B	23.73	15.45

ficant ($F = 4.81$, $df = 2/66$, $p < .025$). The cell means are presented in Table 2.

Trial-by-trial curves for each patient type again showed no overlap between As and Bs, although ABs tended to match (or occasionally exceed) the number of declarative formulations given by the "compatibly" paired groups. ABs thus again did not react differentially to the two patients.

The grammatical form of the "helpful" response is apparently another potent structural index of performance when As and Bs are exposed to the two patient types. Maximum effects were again discerned on the first trial, where the TAS patient elicited declarative formulations from only 3 of 12 As but from 10 of 12 Bs. For the AVOS patient, the respective numbers for As and Bs were eight and two. A 3×2 contingency table for the first trial gave a chi-square value of 7.32 ($df = 2$, $p < .05$). Had one predicted declarative formulations to be emitted in this differential manner, 77% of As and Bs would have been correctly classified on the first trial; 66% of these Ss would still be correctly placed across the five trials.

A-B Status and Interactive Characteristics of Response

Separate 3×2 analyses of variance were performed on the mean number of responses placing into each of the four major Bales categories in five trials. The cell means, shown in Table 3, indicate that regardless of therapist or patient type, relatively more responses placed within the Answers and Questions categories than in the Positive or Negative

TABLE 2
MEAN NUMBER OF DECLARATIVE FORMULATIONS
EMITTED IN FIVE TRIALS

Therapist type	Patient type	
	TAS	AVOS
A	2.33	3.50
AB	3.83	3.75
B	4.08	2.08

(emotional-social) Reaction categories, suggesting that the experimental situation elicited predominantly task-oriented, emotionally neutral responses from Ss.

When responses coded as Positive or Negative Reactions occurred, however, significant Therapist \times Patient interaction effects emerged (Positive Reactions: $F = 4.51$, $df = 2/66$, $p < .025$; Negative Reactions: $F = 3.30$, $df = 2/66$, $p < .05$). These effects, which indicate that the judged emotional quality of "helpful" responses was more positive and less negative under "compatible" pairing conditions, significantly extend the findings relative to length and form of response. Since the increased emission of positively toned responses under compatible pairing conditions appears to have face-valid implications for the "effectiveness" of responses (or at least for rapport-enhancing reactions to patients), the Bales categories add important information to the prior structural indexes of differential performance.

It can be seen that ABs here place between As and Bs in the "emotional" categories and

TABLE 3
MEAN NUMBER OF RESPONSES CODED INTO EACH BALES CATEGORY IN FIVE TRIALS

Therapist type	Bales categories							
	Positive Reactions		Answers		Questions		Negative Reactions	
	TAS	AVOS	TAS	AVOS	TAS	AVOS	TAS	AVOS
A	0.50	1.33	1.42	1.33	2.08	1.75	1.00	0.58
AB	0.92	1.00	2.75	1.92	1.00	1.00	0.33	1.08
B	1.50	0.50	2.25	1.33	1.00	1.67	0.25	1.50

Note.—For each therapist-by-patient combination, the row means sum to 5.0, because each response was coded into one and only one Bales category.

appear to react similarly to the patients. However, ABs show a borderline ($p < .10$) tendency to emit more Answers, regardless of patient type, than the other groups. Mild patient-type main effects were observed in two instances: the AVOS patient tended to elicit more Negative Reactions but fewer Answers than the TAS patient (both effects, $p < .10$).

A-B Status and Content of Response

The three content scales selected, relevance, acceptance, and assertiveness, were applied to the responses given on the first trial only, since it was here that maximal differences between As and Bs were obtained in the analyses for length and form of response. It was hoped that these ratings would clarify the content dimensions involved in the interaction effects.

Table 4 presents the breakdown of the content scale means by therapist and patient type. Analyses of variance showed no main or interaction effects on the acceptance scale; interaction effects on the other two scales, however, were highly significant (assertiveness: $F = 9.24$, $df = 2/66$, $p < .001$; relevance: $F = 6.20$, $df = 2/66$, $p < .005$). Thus, when As responded to the AVOS and Bs responded to the TAS patient, they were rated as more assertive but less relevant than As and Bs paired oppositely. ABs comprised the most assertive and least relevant group of the three, in apparent agreement with their greater reliance on Answers reported above. (Main effects for therapist type on assertiveness, $F = 3.57$, $df = 2/66$, $p < .05$; on relevance, $F = 3.44$, $df = 2/66$, $p < .05$.)

TABLE 4

MEANS OF FIRST-TRIAL CONTENT RATINGS

Therapist type	Content scales					
	Relevance		Acceptance		Assertiveness	
	TAS	AVOS	TAS	AVOS	TAS	AVOS
A	3.25	2.48	3.10	3.29	3.02	3.58
AB	2.33	2.42	3.08	3.48	3.56	3.60
B	2.48	3.14	2.71	3.10	3.63	2.40

Note.—S = relevant, accepting, assertive.

Correlational analyses, in which these content ratings were correlated with five-trial means on the other dependent variables used in this study, clarified the "meaning" of the two content scales on which significant interaction effects were observed. Ratings of assertiveness and relevance appeared most closely related to the dichotomous coding for declarative formulations, such that the emission of declarative formulations was regarded as assertive but not particularly relevant, while (Bales) questions were seen by judges as relevant but not particularly assertive. The third content scale, Acceptance, related positively to Bales Positive Reactions and negatively to Bales Negative Reactions, as one would expect if there were convergence between independent judges' ratings of the "accepting" qualities of these responses. Apart from the acceptance ratings, then, the ratings of first-trial content appear strongly related to the "structural" dependent variables used in the preceding analyses. Ratings of assertiveness, in particular, related to length of response ($r = .35$), declarative formulations ($r = .53$), Bales Positive Reactions ($r = .32$), Bales Answers ($r = .30$), Bales Questions ($r = -.40$), and Bales Negative Reactions ($r = -.11$). All coefficients, excepting the last, were significant at $p < .01$.

Relationships among Measures

To clarify the relationships among those dependent measures which were applied to all five trials, that is, response length, declarative formulations, and the Bales categories, an intercorrelation matrix (Table 5) was computed. The matrix also includes these Ss' subjective reactions (Berzins & Seidman, 1968), to see whether any structural variables analyzed in this study might have acted as "determinants" of postexperimental subjective reactions (satisfaction with responses and ease of selecting helpful responses). Inspection reveals that Ss' subjective reactions are but modestly predictable from the dependent measures used here. Response length is the only variable significantly associated with both subjective reactions.

Response length is also significantly correlated with every other variable in the matrix, suggesting that a general "productivity" or

TABLE 5

INTERCORRELATIONS AMONG DEPENDENT MEASURES AND SUBJECTIVE REACTIONS

Variable	2	3	4	5	6	7	8
1. Response length	51**	27*	39**	-28*	-34**	27*	32**
2. Declarative formulations		47**	72**	-78**	-21	09	27*
3. Bales Positive Reactions			-12	-41**	-16	03	11
4. Bales Answers				-60**	-31**	09	18
5. Bales Questions					-31**	-07	-26*
6. Bales Negative Reactions						01	04
7. Satisfaction with responses							46**
8. Ease of responding helpfully ^a							

Note.—All coefficients are Pearson r 's; decimals have been omitted. $N = 72$.^a Signs for this item, which was scored in the "difficult" direction, have been reflected.* $p < .05$.** $p < .01$.

"responsivity" variable may hold the key to the observed relationships. Substantially positive relationships exist between the "positive" indexes of performance (response length, declarative formulations, Bales Positive Reactions) but their magnitude hardly suggests equivalence. The dichotomization of responses into declarative formulations versus others, however, appears to be almost equivalent to the Bales distinction between Answers and Questions. It is also apparent that Questions and Negative Reactions are uniformly negatively related to other dependent measures.

Further Results

Since, in the original design, each S in this experiment later responded to the "other" patient (subsequent to the responses investigated thus far), the question of the effect of the initial pairings upon later responses may be raised. Briefly, analyses conducted for response length and for declarative formulations indicated that the differential behaviors characterizing S 's performance with the first patient continued unchanged—even though the "other" patient now comprised the stimulus. Trial-by-trial curves continued to show no overlap between As and Bs . While such consistencies in response are intriguing, they must be regarded with caution in the absence of information regarding more precise determinants of the behavioral differences which emerged initially, for example, on the first trial with the first patient.

DISCUSSION

In interpreting the results of this study, it is well to note at the outset that the categories of analysis were not selected beforehand, but, in the absence of appropriate precedents, resulted from a concern with discovering the "natural lines of cleavage," if any, in the "helpful" responses of these untrained Ss .

The major findings of this study were that when As responded "helpfully" to the schizoid patient, and Bs responded to the neurotic patient—relative to the opposite pairings—they emitted responses characterized by greater length, by declarative rather than interrogatory form, and by positive (social-emotional) overtones. It has already been established that these Ss , subsequent to performing in this manner, felt relatively satisfied with their performance and reported greater ease in selecting helpful responses (Berzins & Seidman, 1968). Persons intermediate in A-B status (ABs) did not react differentially to the two patients. The differences in the therapeutic responses of As and Bs , which were marked after as little as 1 minute of exposure to the "patients," moreover tended to persist across trials; at no point was overlap between As and Bs observed.

For these untrained Ss , the somewhat correlated measures of length and form of response appeared to account for the major variations in emitted "therapeutic" responses.

A relatively parsimonious integration of the results can be evolved around an "information-processing" model. That is, when As and Bs were compatibly paired with patients (i.e., As with AVOS and Bs with TAS), they behaved as though they had assimilated and processed enough information to be able to produce a declarative, positively toned response of appreciable length (which Ss felt to be satisfying and helpful). When As and Bs were incompatibly paired (As with TAS and Bs with AVOS), they behaved as though they had not obtained enough information to be able to formulate a helpful response; they, in effect, "requested more information" from the patient (engaged in shorter, negatively toned inquiries which they felt to be less satisfying and less helpful). Subsequent patient communications, which in fact provided a great deal of relevant "information," failed to modify the responses of compatibly or incompatibly paired Ss; declarers continued to declare and questioners to question. The initial pairing conditions thus appear to determine how much information Ss are able to "process" in the interest of formulating helpful responses.

What features of the patient communications can be held responsible for the observed differences in response? The tape-recorded communications were designed to expose Ss to the TAS and AVOS symptom clusters in a naturalistic manner; there was no attempt to equate the tapes for dramatic content or "pull" at any one point. In scrutinizing the first segments for decisive contrasts, one clear difference emerges in the dysphoric affect of the TAS patient (which is not present in the blander, faster speaking AVOS patient) and the cognitive disorganization (disjointed speech, vagueness) of the AVOS patient. Perhaps the disjointed utterances of the AVOS patient are particularly troublesome to Bs in light of their reportedly greater need for cognitive clarity and field-independent perceptual-cognitive style (Pollack & Kiev, 1963; Shows & Carson, 1965) or perhaps Bs themselves would tend to rely upon AVOS-like adjustable techniques under stress (Berzins, Friedman, & Seidman, 1969; Sandler, 1965). As in turn may be disrupted by the TAS patient's dysphoria and

intropunitiveness, since several studies suggest that As themselves may display depressive, intropunitive behaviors under stress (Berzins et al., 1969; Sandler, 1965). This line of reasoning suggests that the emission of helpful responses is impeded by Ss' "blind spots" if these happen to resemble those of the patient. If this notion has merit, then the present results would argue for the efficacy of pairing therapists with patients on the basis of dissimilar reactions to stress, which may be partially indexed by a complementary (not similar) A-B status. In any case, one important heuristic implication of the present research is that since marked effects were observed on the very first trial, a systematic search for the precise cues involved in triggering the differential "helpful" behaviors can be undertaken.

In considering the generality of the effects observed in this study, considerable caution is in order, however. While it may seem that the structural indexes of differential performance (especially length of response) coupled with Ss' own subjective reactions to the experience, can be interpreted as personality-based antecedents to differential performance in psychotherapy, it must be stressed that, unlike in actual therapy, Ss in this study received no feedback from the patients which in any way could modify their subsequent performance. The absence of trial-to-trial overlap between As and Bs (which continued unchanged when the "other" patient was introduced) must probably be interpreted as stereotyped responding in untrained Ss. The effects of first-trial pairings were substantial, but whether such effects apply to the performances of practicing therapists is a matter for further research. It would be desirable to obtain evidence on the treatment performances of AB therapists to facilitate comparisons with As and Bs, since ABs did not consistently occupy an intermediate position on the dependent measures used in this study. In this way, important information regarding the possible utilization of mental health volunteers and the feasibility of therapist-patient matching could be obtained.

On the whole, the results of the present experiment suggest that the A-B variable is intrinsically related to variations in "helpful"

performance in male dyads. While the reasons for the predictive power of the A-B variable are only beginning to be understood, the differential responsiveness of As and Bs to varying forms of disordered behavior provides an important variance-reducing ingredient for the design of psychotherapy studies.

REFERENCES

BALES, R. F. *Interaction process analysis*. Cambridge, Mass.: Addison-Wesley, 1950.

BERZINS, J. I., FRIEDMAN, W. H., & SEIDMAN, E. Relationship of the A-B variable to patient symptomatology and psychotherapy expectancies. *Journal of Abnormal Psychology*, 1969, 74, 119-125.

BERZINS, J. I., & SEIDMAN, E. Subjective reactions of A and B quasi-therapists to schizoid and neurotic communications: A replication and extension. *Journal of Consulting and Clinical Psychology*, 1968, 32, 342-347.

BETZ, B. J. Studies of the therapist's role in the treatment of the schizophrenic patient. *American Journal of Psychiatry*, 1967, 123, 963-971.

CARSON, R. C. A and B therapist "types": A possible critical variable in psychotherapy. *Journal of Nervous and Mental Disease*, 1967, 144, 47-54.

CARSON, R. C., HARDEN, J. A., & SHOWS, W. D. A-B distinction and behavior in quasi-therapeutic situations. *Journal of Consulting Psychology*, 1964, 28, 426-433.

McNAIR, D. M., CALLAHAN, D. M., & LOER, M. Therapist "type" and patient response to psychotherapy. *Journal of Consulting Psychology*, 1962, 26, 425-429.

PHILLIPS, L., & RABINOVITCH, M. Social role and patterns of symptomatic behavior. *Journal of Abnormal and Social Psychology*, 1958, 57, 181-186.

POLLACK, I. W., & KIEV, A. Spatial orientation and psychotherapy: An experimental study of perception. *Journal of Nervous and Mental Disease*, 1963, 137, 93-97.

ROGERS, C. R., GENDLIN, E. T., KIESLER, D. J., & TRUAX, C. B. *The therapeutic relationship and its impact*. Madison: University of Wisconsin Press, 1967.

SANDLER, D. *Investigation of a scale of therapeutic effectiveness: Trust and suspicion in an experimentally induced situation*. (Doctoral dissertation, Duke University) Ann Arbor, Mich.: University Microfilms, 1965. No. 66-1382.

SHOWS, W. D., & CARSON, R. C. The A-B therapist "type" distinction and spatial orientation: Replication and extension. *Journal of Nervous and Mental Disease*, 1965, 141, 456-462.

STRUFF, H. H. A multi-dimensional system for analyzing psychotherapeutic techniques. *Psychiatry*, 1957, 20, 293-306.

WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

(Received June 3, 1968)

EXPERIMENTAL APPROACH TO DEFINING THE ROLE OF SOCIAL DESIRABILITY IN PERSONALITY ASSESSMENT:

IS THERE ONE RESPONSE PROCESS OR TWO? ¹

STEPHEN LEFF ² AND NANCY LAMB

Washington University

This study was designed to determine whether Ss engage in one or two different response processes when they endorse items as descriptive of their selves and their personal ideals, and when they rate the items' characteristic social desirability. The experiment used three sets of 2×2 factorial designs, one for each rating category (self, personal ideal, and social ideal). The Ss' need for social approval and the instructional set were the variables controlled. The analyses of variance showed that Ss do engage in two very different kinds of responding when they are asked to describe their selves and their personal ideals, and when they are asked to describe the socially desirable. The first two rating categories were differentially influenced by the independent variables. In contrast, the finding that neither of these variables was instrumental in determining the nature of the response in the third category lends support to the hypothesis that items do have a characteristic social desirability which is derived from their perceived cultural value.

Edwards (1957), Kenny (1956), and others have demonstrated that in structured psychometric situations, the probability that an individual will endorse a particular item as descriptive of his self or his ideal can be predicted accurately by the location of that item on a continuum derived from group judgments of item social desirability. Edwards (1957) has hypothesized that this phenomena reflects the tendency of Ss to implicitly judge and respond to inventory items in terms of their characteristic cultural value even when overtly engaged in self-descriptive or ideal-descriptive tasks.

Messick (1960), Crowne and Stephens (1961), and Scott (1963), on the other hand, have hypothesized that these high correlations are determined by multiple response sets which mirror not merely the characteristic cultural value of items, but also the need of Ss to conform or to respond in a socially

approved fashion, and other situational and interpersonal variables. As evidence for this hypothesis Messick (1960) cites his findings that the social desirable dimension is composed of nine identifiable factors, only one of which appeared to reflect the social desirability variable as Edwards defined it.

The hypotheses of Messick and the others who share his theoretical position are based on the assumption that Ss engage in at least two very different processes when they endorse items as descriptive and when they rate the characteristic social desirability of the items, processes that may be related but which are by no means identical. In contrast, Edwards' hypothesis is based on the assumption that when Ss endorse inventory items as descriptive of their selves or ideals they are engaging in essentially the same process that they follow when they make judgments about the social desirability of those items.

This study attempted to further delineate the positions of Edwards and Messick with the possibility of furnishing support for either of them. The Ss with differing personality characteristics were asked to describe their selves, their personal ideals, and the socially desirable, while at the same time their understanding of the Es' hypothesis about the experiment was manipulated. It was reasoned

¹ This paper was presented as part of a symposium at the meeting of the Southwestern Psychological Association, Houston, April 1967. The authors wish to express their appreciation to Malcolm D. Gynther for his critical comments in the preparation of this manuscript.

² Now at Harvard Medical School. Requests for reprints should be sent to Stephen Leff, Massachusetts Mental Health Center, 74 Fenwood Road, Boston, Massachusetts 02115.

that if differing response processes were, in fact, involved in the separate tasks, the responses in each descriptive category would be differentially affected by the personality and/or situational variables manipulated. Conversely, it was predicted that if a common response process underlay the modes of responding, the responses in each category would not be differentiated from each other with respect to the effects of the independent variables.

METHOD

Subjects

Sixty white male undergraduates ranging in age from 18-20 were drawn from a sample of 71 original volunteers. (The explanation of this reduction in sample size is presented in procedural considerations.)

Instruments

The instruments used included the Crowne-Marlowe test of the need for social approval (Crowne & Marlowe, 1964) and the Interpersonal Check List for descriptions of personality traits (Leary, 1957).

Procedure

In order to ascertain their attitudes toward the instructions which they were to be given, Ss were asked whether they thought good adjustment was characterized by having high ideals and living up to them (idealistic attitude) or by having moderate ideals and recognizing the areas in which one fails to live up to them (realistic attitude). The *Es* originally contemplated having equal numbers of "idealistic" and "realistic" Ss; however, the original *S* population was such that the "realistic" attitude was chosen by 69 of the original 71 Ss. Therefore, it was not possible to contrast what might be considered personal set or orientation, and the two Ss who chose the "idealistic" alternative were dropped from the study. The sample size was then reduced from 69 to 60 by randomly eliminating Ss from each of the experimental groups with more than 15 Ss.

Each *S* completed the Crowne-Marlowe test and was instructed as to the *Es*' hypothesis: one-half of the Ss were informed that good adjustment was hypothesized to be characterized by having high ideals and living up to them, and the other half was told that good adjustment was hypothesized to be a function of having moderate ideals and of recognizing the ways in which one failed to live up to them.

Following these instructions, the Interpersonal Check List was administered in booklet form and Ss were asked to complete the booklet for the following categories: For the first category (self) Ss were asked to describe the kind of person they were at that time; for the second (personal ideal)

they were asked to describe the kind of person they would like to be; for the third (social ideal) Ss were asked to describe the person they thought society found most desirable.

For the purpose of data analysis this experiment used three 2×2 factorial designs, one for each of the descriptions Ss were asked to make (self, personal ideal, social ideal). Instructional set (either "idealistic" or "realistic") and Ss' need for social approval (either "high" or "low" as determined by dividing the Crowne-Marlowe scores at the median) were the independent variables.

Item clusters (Leary's octant scores) rather than single items were used as dependent variables in this study to increase reliability and to facilitate the analysis of the processes under investigation. Dominance-submission (Dom) and love-hostility (Lov) summary scores were also computed in the manner described by Leary, transformed into standard scores, and converted to single summary points on Leary's two-dimensional grid. Difference scores among the three summary points of the descriptive categories (self, personal ideal, social ideal) were then calculated for each *S*. These difference scores, along with the summary scores, were analyzed in a supplementary 2×2 analysis of variance.

RESULTS

Tables 1 and 2 summarize the analyses of variance of Ss' self, personal ideal, and social ideal descriptions as a function of need for social approval and instructional set. These tables reflect the fact that Ss high in the need for social approval checked significantly more adjectives as descriptive of themselves in octants measuring affiliation and responsibility (7 and 8) and obtained higher Lov scores than did those low in that need. These high need for social approval Ss similarly described their personal ideals by checking more adjectives than low Ss in Octants 7 and 8, and in Octant 6 (Respectful). Table 2 indicates that they also achieved higher Lov scores for the personal ideal category.

A further inspection of Tables 1 and 2 reveals that although the instructional set had a limited influence on Ss' self-descriptions, there was no marked or specific pattern. Instead, it effected primarily the way Ss described their personal ideals. On this category the instructions influenced those octants measuring power or lack of power (1 and 5), independent strivings (2), and rebelliousness (4). In all cases Ss who had been told that good adjustment is a function of high ideals

TABLE 1
SUMMARY OF THE ANALYSES OF VARIANCE OF SUBJECT'S SELF, PERSONAL IDEAL, AND SOCIAL IDEAL OCTANT SCORES
AS A FUNCTION OF NEED FOR APPROVAL AND INSTRUCTIONAL SET

Source	Octant 1			Octant 2			Octant 3			Octant 4			Octant 5			Octant 6			Octant 7			Octant 8		
	df	MS	F	df	MS	F	df	MS	F	df	MS	F	df	MS	F	df	MS	F	df	MS	F	df	MS	F
Self-description	1	8.07	0.90	1	2.40	0.38	1	5.40	0.94	1	15.00	2.68	1	5.40	0.67	1	4.27	0.94	1	45.07	11.65****	1	36.82	6.85***
Approval (A)	1	19.27	2.12	1	13.07	2.60	1	4.27	0.74	1	1.40	0.43	1	0.60	0.08	1	3.27	0.72	1	6.67	1.72	1	20.42	3.80*
Instruction (B)	1	0.60	0.07	1	9.60	1.52	1	26.67	4.65**	1	0.27	0.05	1	2.40	0.30	1	2.40	0.52	1	9.60	2.48	1	8.81	1.64
A X B	56	9.10		56	6.30		56	5.73		56	5.62		56	8.01		56	4.60		56	3.86		56	5.36	
Error																								
Personal ideal	1	0.07	0.03	1	1.67	0.73	1	0.27	0.09	1	0.07	0.04	1	0.00	4.39**	1	22.82	6.54****	1	25.35	5.48***	1	19.27	3.71*
Approval (A)	1	8.07	3.90*	1	8.07	3.54*	1	8.07	2.74	1	13.07	6.74***	1	5.40	4.39**	1	4.82	1.38	1	2.82	0.60	1	6.67	1.28
Instruction (B)	1	8.07	3.90*	1	4.27	1.87	1	0.07	0.02	1	0.27	0.14	1	0.27	0.22	1	0.02	0.01	1	1.35	0.29	1	9.60	1.85
A X B	56	2.07		56	2.28		56	2.95		56	1.94		56	1.23		56	3.49		56	4.62		56	5.20	
Error																								
Social ideal	1	0.42	0.39	1	3.27	1.63	1	0.02	0.01	1	7.35	3.91*	1	0.27	0.18	1	6.02	1.33	1	3.27	0.40	1	5.40	0.87
Approval (A)	1	0.83	0.75	1	2.40	1.19	1	0.82	0.82	1	0.02	0.01	1	6.67	3.94*	1	7.35	1.62	1	8.07	0.99	1	6.67	1.07
Instruction (B)	1	0.15	0.14	1	4.27	2.12	1	0.42	0.21	1	2.02	1.07	1	0.27	0.175	1	0.02	0.00	1	24.07	2.94*	1	1.07	0.17
A X B	56	1.09		56	2.01		56	1.96		56	1.88		56	1.54		56	4.54		56	8.18		56	6.24	
Error																								

Note.—The meanings for the eight octants are as follows: Octant 1: Managerial-Autocratic; Octant 2: Competitive-Narcissistic; Octant 3: Aggressive-Sadistic; Octant 4: Rebellious-Distasteful; Octant 5: Self-Effacing-Masochistic; Octant 6: Docile-Dependent; Octant 7: Cooperative-Overconventional; Octant 8: Responsible-Hypernormal.

* $p < .10$.
 ** $p < .05$.
 *** $p < .025$.
 **** $p < .001$.

TABLE 2

SUMMARY OF THE ANALYSES OF VARIANCE OF SUBJECT'S
SELF, PERSONAL IDEAL, AND SOCIAL IDEAL DOM
AND LOV SCORES AS A FUNCTION OF NEED
FOR APPROVAL AND INSTRUCTIONAL SET

Source	DOM			LOV		
	df	MS	F	df	MS	F
Self-description						
Approval (A)	1	244.02	2.82*	1	340.82	7.19***
Instruction (B)	1	138.02	1.59	1	8.82	0.19
A \times B	1	2.81	0.03	1	268.81	5.68**
Error	56	86.61		56	47.35	
Personal ideal						
Approval (A)	1	6.02	0.74	1	209.07	6.81**
Instruction (B)	1	0.15	0.02	1	4.27	0.14
A \times B	1	18.15	2.22	1	32.26	1.05
Error	56	8.18		56	30.72	
Social ideal						
Approval (A)	1	0.60	0.05	1	86.40	1.99
Instruction (B)	1	6.67	0.53	1	26.67	0.61
A \times B	1	0.00	0.00	1	1.07	0.02
Error	56	12.48		56	43.42	

* $p < .10$.

** $p < .025$.

*** $p < .001$.

checked significantly more of the items in these categories than did those who had been told that the moderate ideals reflected good adjustment.

Finally, as Table 1 indicates, the category of social ideal manifested no definitive trends as a function of either the need for social approval or instructional set. In no instance did either of the independent variables yield significant differences in more than one octant. Thus, although significant differences at the .10 level were obtained as a function of the need for approval in Octant 4 and as a function of instructional set in Octant 5, as single statistics they do not appear to be parts of any meaningful pattern. This interpretation is supported by the fact that, as Table 2 indicates, the independent variables had no significant effects on either the Dom or Lov summary scores.

A supplementary analysis of variance of the difference scores among Leary's summary points indicated trends approaching significance ($p < .10$) for Ss high in the need for social approval to perceive their self-ratings as being closer to their personal and their social ideals than did Ss low in the need for social approval. The Ss who had received idealistic instructions described their personal ideals as being significantly more different from their social ideals ($p < .05$) than did

Ss who received the realistically oriented instructions.

CONCLUSIONS AND IMPLICATIONS

The results of this study support Messick (1960) and those others who hypothesize that when Ss judge the social desirability of inventory items (social ideal) they engage in a very different kind of response process than they do when asked to describe their selves or their personal ideals. When Ss gave self-descriptions, the need for social approval (which is a discriminable personality characteristic) affected their scores, but instructional set had little specific influence. When Ss were asked to describe their personal ideals they responded partly on the basis of this personality variable and partly on the basis of the instructional set they were given. When asked to describe the social ideal, however, neither of these factors, instructional set or need for approval, was instrumental in determining the nature of the responses.

It is also true, however, that this study also lends support to Edwards' (1957) hypothesis that items do have a characteristic social desirability which is derived from their perceived cultural value and which is not significantly affected by some situational and personality variables. In spite of differing needs for social approval and differing instructional sets, Ss describing their social ideal agreed for almost all octants on the degree to which each was socially desirable.

The finding that Ss' personal ideals were most susceptible to the influence of situational variables than either self-ratings or social ideal ratings suggests that instructions to rate one's ideal also elicit a distinctive response set. One suggestive explanation of this fact might be that when Ss describe their personal ideals they are, by the nature of the task, neither bound by the actuality of their own behavior nor by their perceptions of the cultural norms, and in the absence of these guidelines they become more susceptible to the influence of *E*. One implication of this finding would appear to be that the influence of an *E* on the responses of his Ss is mediated by the nature of the dependent variables he chooses to include in his study.

The results of the difference scores analyses are consistent with the previous studies of Kenny (1956), which have shown that Ss high in the need for social approval saw their actual behavior as less discrepant from either their personal ideals or the socially desirable than did those low in this need. A finding not so predictable, which would seem to warrant further research, was that Ss who were informed that good adjustment is a function of having high ideals and living up to them tended to describe their personal ideals and social ideals as being more discrepant than those who received the other set of instructions. Perhaps these Ss took their instructions as stressing an "individualistic" orientation and counterindicating social conformity.

REFERENCES

- CROWNE, D., & MARLOWE, D. *The approval motive*. New York: Wiley, 1964.
- CROWNE, D., & STEPHENS, M. Self-acceptance and self-evaluative behavior. *Psychological Bulletin*, 1961, **58**, 104-121.
- EDWARDS, A. *The social desirability variable in personality assessment and research*. New York: Dryden Press, 1957.
- KENNY, D. T. The influence of social desirability on discrepancy measures between real self and ideal self. *Journal of Consulting Psychology*, 1956, **20**, 315-318.
- LEARY, T. *Interpersonal diagnosis of personality*. New York: Ronald Press, 1957.
- MESSICK, S. Dimension of desirability. *Journal of Consulting Psychology*, 1960, **24**, 279.
- SCOTT, W. Social desirability and individual conceptions of the desirable. *Journal of Abnormal and Social Psychology*, 1963, **67**, 574-585.

(Received June 5, 1968)

ATTITUDES TOWARD DEATH:

A PSYCHOLOGICAL PERSPECTIVE¹

HERMAN FEIFEL²

Veterans Administration Outpatient Clinic, Los Angeles

As a consequence of psychology's enhanced interest in the existential richness of human behavior, research in the area of death is becoming a productive enterprise. The influencing aspect of consciousness of death is active over the entire age continuum and is not restricted to the sick, aged, suicidal, or combat soldier. The meaning of death is multidimensional and varies not only between individuals but within the same person. Additionally, dying and death possess psychological and social features and are not merely biological events. Our overall socially repressive orientation toward death promotes neurotic concerns about it. More intimate familiarity with death is required in our developmental upbringing and culture.

Two major developments influenced psychology's halting approach in probing attitudes toward death as a relevant variable in human conduct. One was psychology's natal need to raise its flag independently of philosophy and ethics. Scientific respectability meant occupying oneself with measurable stimuli and responses which were repeatable and public. A consequence was restraint in scrutinizing consciousness as a datum and neglect of personality. The other was a waning of faith in the providential and sacred accompanied by a lack of transcending significance for death. This secular turn was further associated with a decline in the kinship group and increasing social fragmentation of the family, a growing impersonality emerging from a technologically dominated society, and deritualization of grief and mourning practices—all fusing to exacerbate our hostility toward and repudiation of death.

The events of World War II, however, and the ensuing press of urgent social problems forced psychology to look beyond its traditional positivism and enter into the bailiwick of moral issues and life choices. The view

that a vital psychology must be rooted in man, not in a mathematical physics model, became more insistent. At the same time, a legacy of the same war thrust all of us into a baleful situation in which the world annihilation fantasies of the schizophrenic patient no longer appeared unrealistic. Physical science has presently made it possible for us all to share a common epitaph. Not only the individuality of death but posterity and social immortality are now menaced.

In this junction of psychology's waxing interest in the existential richness of human life, and sharpened awareness of life's temporality and recognition that all truth is germane to the exigencies of survival, research in the area of death is emerging as an authentic and fertile enterprise.

Consideration of death is undoubtedly one of the foremost sources of anxiety for man. Indeed, a major task of both religion and philosophy has been to grapple with its intensity and complexity. Psychology, in coming to grips with the meaning of death for the individual, asserts its responsibility to reexamine its philosophic and humanistic heritage in the context of its generic ideal of science.

Investigation of attitudes toward death and bereavement makes available a supplementary entryway in understanding adaptational strategies used in coping with pain, crisis, and stress. It further affords the possibility of improving treatment of the seriously sick and others facing the prospect of oncoming death. Additionally, it extends our comprehension of

¹ A version of this paper was presented at the University of Minnesota symposium "Death, Grief and Bereavement: An Interdisciplinary Evaluation," Minneapolis, May 18, 1967. Preparation of the paper was facilitated by National Institute of Mental Health Grant M-2920 and the Veterans Administration Medical Research Program.

² Requests for reprints should be sent to Herman Feifel, Veterans Administration, 1031 S. Broadway, Los Angeles, California 90015.

how death affects the social fabric of society. Overriding these, however, may be our due recognition of the importance of the future in steering present behavior. It is man's excellent capacity to conceptualize a future—and inevitable death—which distinguishes him from other species and makes myopic imitative application of a physical sciences methodology to him. Anticipation guides a majority of our actions and expectation serves as a principal mediator of goal-directed and purposeful behavior. As human beings, the present contains not only the psychological past but psychological future as well. Death is an absent presence even before its actual arrival. Only with consciousness of mortality does concern with human fulfillment, salvation, and the notion of the uniqueness and individuality of each one of us receive complete meaning.

The piloting force of awareness of death is active at all age levels. Its domain is not confined to the sick, aged, suicidal, or combat soldier. In this regard, it may be pertinent to emphasize that we not shut out children from the realities of death. The efforts of investigators as Anthony (1940), Maurer (1966), Nagy (1948), and Wahl (1958) disclose that we do not protect the child by attempting to shield him from death. Rather, we hinder his emotional growth. In actuality, this proscription displays more of the adult's or parent's own anxieties and strong emotions concerning death than the child's genuine capacity to manage knowledge of the existence of death. Naturally, level of explanation, timing, and individual differences have to be appreciated.

Until recently, the psychoanalytic conceptualization of attitudes toward and fear of death as derivative events was dominant. Sentiments about death were essentially manifestations of a more ultimate reality, that is, separation anxiety or conflicts about castration. Doubtless, such clinical displacement does occur. Death fears can be secondary phenomena. Nevertheless, incoming data increasingly suggest that the reverse may be more to the point. Apprehensiveness over bodily annihilation and concerns about finitude themselves assume dissembling guises. The depressed mood, fears of loss, sundry psychosomatic symptoms, and varying psy-

chological disturbances all have evidenced affinity to anxieties concerning death. Gillespie (1963), the English psychiatrist, for example, goes so far as to declare that fear of death is the most traumatic factor in fashioning senile psychosis. Feifel (1959) and Searles (1961) have both advanced the hypothesis that *one* of the intentions of schizophrenic denial of reality, in certain persons, may function as a magical holding back, if not undoing, of possible death. If living leads inevitably to death, the way not to die is not to live. We need to incorporate thoughts and emotions about death into our personality studies and therapeutic sights. Too frequently do we number their impetus by engaging ourselves with the symbolic as divorced from the *real* aspects of death.

We are currently more knowing about the multidimensional meanings which death holds for people. Death can be the "gentle night" or "great destroyer." It can reflect surcease from pain and tribulation, reunion with one's family, loss of control, punishment, or loneliness. Developmental changes, cultural conditioning, religious orientation, personality characteristics, and level of threat, among others, contribute their shaping power.

There is also the implication that attitudes toward death can oscillate in the same individual, from strong avoidance through anxious hope and uneasy resignation to calm acceptance. The point in time one centers on can be crucial. In this context, Feifel (1965) has indicated the coexistence of contradictory attitudes toward death, for example, realistic acceptance of death and its simultaneous rejection in a subtle equilibristic balance in many seriously ill and terminally ill patients. He has suggested that this counterpoise seems to serve adaptational needs of the patient, allowing him to maintain communal associations and yet organize his resources to contend with oncoming extinction.

Additionally, we are discovering that treatment of the seriously ill and dying may be too conditioned to considerations and wants of the healthy rather than to the needs of the sick. Research findings underscore the desire of most patients for honest and sharing talk from physicians and family about the seriousness of their illness (Feifel, 1959). This is in

strong contrast to the prominent orientation of physician, nurse, and family to avoid such confronting discussion (Feifel, 1965; Glaser & Strauss, 1965). The threat to professional narcissism and existing social scruples about death are obviously contributory. Reappraisal of what and how we instruct our medical, nursing, and seminary students concerning dying, death, grief, and bereavement seems in order.

Truth and what sustains one's sense of reality tend to be more supporting and beneficial to the patient than deception and denial. We must resist disinheriting the patient psychologically and socially at the very time we enhance attention to his physiological requirements. Approaching death is better integrated when the patient feels understood and can voice his feelings and thoughts about dying and death. Clinical experience demonstrates that when this happens, less depression, deviant behavior, and blame of others show themselves in patients. Further opening channels of communication tends to attenuate feelings of inadequacy and guilt not only in the patient but in the professional people involved in treatment and family as well.

Some terminally ill patients, when apprised of their prognosis, manifest an adjustment equal to or better than that seen in similar patients who have not been so informed (Zinker, 1966; Zinker & Fink, 1966). Realization of possible death appears to give permission, even freedom, to some to be themselves rather than extensions or mirrors of other people's values. One can learn, apparently, not only to live with enormous stress but to grow and change under its goad. Indeed, history furnishes us with numerous examples where recognition of impending death has fostered creativity rather than paralysis.

Magnified cognition exists that death represents a fact of social as well as psychological import. The funeral ceremony, although personal in design, is certainly societal in consequence. A person's death affects the emotional life of the bereaved but it also entails changes in the status and social position of the survivors.

What are some deductions from these perceptions? Man is a creature in time and space

whose consciousness permits him to nullify their strictures. This brings into his ken the awareness of inexorable death as a compelling stimulus. As a time-ridden being, man is faced with the task of identifying himself with eternity. His most viable response will issue from basic philosophical, religious, or psychological deliberations about death already in his possession. The agony of selfhood is not endurable for most of us in pondering death without resources, be they transcendental, inspirational, or existential. We pay expensively for the taboo we affix to the subject of death. Closer psychological familiarity with death is called for in our developmental upbringing and culture. Our socially repressive outlook encourages neurotic anxieties about death. The evolution of an *ars moriendi* prior to the advent of death is the charge.

Cutting across all the populations the author has studied (Feifel, 1959, 1965) is the view of death as the extreme abomination in man's experience. Despite invocation of denial and a host of other stratagems, fear of finitude is never quite hushed. Withstanding all jeremiads and alienation, in the face of death the vast majority in all the author's groups, that is, terminally, seriously, chronically, and mentally ill, as well as normal subjects, assess the world as good and not to be ceded if possible.

Man has a legitimate need to face away from death. In truth, who is to say that under certain conditions this may not be salutary? Unfortunately, excessive camouflage and expulsion of the notion of death prevail in the United States leading to a falsification of the essence of man. If we accept death as a necessity rather than strive to demote it to the level of accident, there may be less need to project fear of death outside ourselves. This might possibly mute some of the violence of our times. Energies now bound up in continuing attempts to shelve and repress the concept of death would be available to us for more constructive aspects of living—perhaps even fortify man's gift for creative splendor against his genius for destruction. Forster (1962, p. 394) has commented penetratingly that "death destroys a man but the *idea* of death saves him."

Finally, what of psychology's more specific

assignment in the area of death? Psychotherapeutic functioning and models of personality and psychopathology need amplified representation of the future and death in their horizons. Findings suggest that one's philosophy of life and death lies at the nexus of meaning, value, and personality.

The thrust of empirical attack requires extension. Theoretical expressions must be more rooted in the reasonably precise bases of systematic observation and experimentation. It may be that psychology is particularly favored to be enlightening here since some feel that it is a sector where humanistic and physicist-engineer cultures, too often insulated and at cross-purposes with one another, intersect. Clearly, psychology's own imperative for interdisciplinary perspective should not be minimized.

Methodology will have to capture, in more astute form, the manifold meanings of death not only between individuals but within the same person. We must tighten our grasp of the bonds existing among verbally expressed attitudes toward death, fantasy notions, and below-the-level-of-awareness ideas. As in other areas, we have the bedeviling challenge of making constructs, such as fear of death, operationally definable yet integrating the cogency of such explanatory concepts as purpose, anxiety, redemption.

The pivotal undertaking for psychology has been alluded to by Murphy (1959, p. 320): "Instead of looking at the grim skeleton and crossbones and being told *this* is reality, let us study the observer himself, regard him as

the reality and ask why skulls and crossbones are the chief realities for *him*."

Only by encompassing the concept of death into his life will man fully understand himself—and only the culture which countenances death will truly savor life.

REFERENCES

- ANTHONY, S. *The child's discovery of death*. New York: Harcourt, Brace, 1940.
- FEIFEL, H. *The meaning of death*. New York: McGraw-Hill, 1959.
- FEIFEL, H. The function of attitudes toward death. In, *Death and dying: Attitudes of patient and doctor*. Vol. 5, Symposium No. 11. New York: Group for the Advancement of Psychiatry, 1965.
- FORSTER, E. M. Aphorism. In W. H. Auden & L. Kronenberger (Eds.), *The Viking book of aphorisms*. New York: Viking Press, 1962.
- GILLESPIE, W. H. Some regressive phenomena in old age. *British Journal of Medical Psychology*, 1963, 36, 203-209.
- GLASER, B. G., & STRAUSS, A. L. *Awareness of dying*. Chicago: Aldine, 1965.
- MAURER, A. Maturation of concepts of death. *British Journal of Medical Psychology*, 1966, 39, 35-41.
- MURPHY, G. Discussion. In H. Feifel (Ed.), *The meaning of death*. New York: McGraw-Hill, 1959.
- NAGY, M. H. The child's theories concerning death. *Journal of Genetic Psychology*, 1948, 73, 3-27.
- SEARLES, H. Schizophrenia and the inevitability of death. *Psychiatric Quarterly*, 1961, 35, 631-665.
- WAHL, C. W. The fear of death. *Bulletin of the Menninger Clinic*, 1958, 22, 214-223.
- ZINKER, J. C. *Rosa Lee: Motivation and the crisis of dying*. Lake Erie College Studies, Vol. 6. Painesville, Ohio: Lake Erie College Press, 1966.
- ZINKER, J. C., & FINK, S. L. The possibility for psychological growth in a dying person. *Journal of General Psychology*, 1966, 74, 185-199.

(Received June 7, 1968)

SENSING HUMOR: LATENCY AND AMPLITUDE OF RESPONSE RELATED TO MMPI PROFILES¹

DOUGLAS S. HOLMES²

Stockton State Hospital

Forty-eight male Ss responded to cartoons by (a) sorting one set into four equal piles ranging from Most Funny to Least Funny, and (b) reacting immediately to each cartoon of another set by manipulating a knob. The Ss were divided, according to their MMPI profiles, into *PPd* (highest score on the Psychopathic Deviate scale), and *NPPd* (highest score on some other scale). Consistent with expectations, *PPd* Ss preferred sexual (and hostile) humor more than *NPPd* Ss, preferred sexual and hostile humor more than nonsense humor, and responded to all cartoons more quickly than *NPPd* Ss. The immediate reaction measures of latency and amplitude appear to hold promise for future research on humor.

Lack of systematic knowledge concerning humor, rather than lack of relevance, probably accounts for the infrequency with which humor is discussed in contemporary psychological literature. Only once in the 18 years of publication of the *Annual Review of Psychology* has humor been mentioned. On that occasion, Jenness (1962) commented: "Sense of humor, long respected as a personality variable, has proved elusive as an object of research [p. 506]." Perhaps its elusiveness in the laboratory is a function of our devices of detection rather than a function of the esoteric nature of humor. Allport (1954) implied as much: "There is . . . one . . . trait that, up to now, has defied successful psychological study, viz. *sense of humor* . . . its accurate measurement is beyond the present competence of psychology [p. 437]."

Throughout the years there have been remarkably few variations in the techniques of measuring response to humor. In general, the statement of Flugel (1954) remains valid today: "Experimentation has . . . taken the

form of the presentation of humorous stimuli . . . the subject being asked to rank the items in the order of funniness, or to give them marks in accordance with a predetermined scale [p. 709]."

It seems clear that more sensitive measures of response to humor would be useful to those studying humor. Accordingly, the primary purpose of this study was to try out an innovation in measuring an S's response to humor. Instead of relying totally on traditional, judgmental rating responses of humor preference, measures were obtained which more closely approximate the perceptual process of humor-being-sensed.

In order to provide a context for the innovation in measuring sense of humor, cartoons were categorized into sex, hostility, and nonsense; Ss were divided into two groups, one with psychopathic tendencies and one without such tendencies; and mean cartoon scores were compared between groups. Previous research and theory suggested that personality correlates would be found using this procedure (Dahlstrom & Welsh, 1960; Kambourpovlov, 1926; Luborsky & Cattell, 1947; Tollefson & Cattell, 1963-1964). The present report describes some differences in response to humor between the two groups, discusses implications of the obtained differences, and discusses the potential value of the new measures that were used.

¹This study was supported in part by a grant from the California State Department of Mental Hygiene. The author wishes to express his appreciation to Dianna Minnehan for her assistance in data collection and analysis.

²Requests for reprints should be sent to Douglas S. Holmes, who is now at Smith Richardson Foundation, P. O. Box 3265, Greensboro, North Carolina 27402.

METHOD

TABLE 1

MEAN AGE, VOCABULARY, AND MANIFEST ANXIETY OF *PPd* AND *NPPd* Ss

Item	<i>PPd</i>		<i>NPPd</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age	32.50	11.40	31.11	8.97
Vocabulary	31.55	3.36	31.07	3.97
Manifest anxiety	10.10	4.83	9.46	6.33

Note.—For *PPd*, *N* = 20; for *NPPd*, *N* = 28.

Subjects

Male psychiatric technicians and technician trainees who had worked at Stockton State Hospital for 7 years or less were asked to participate during working hours. Virtually all were willing. Scheduling was feasible for 48 men, all of whom were high school graduates.

Conditions

The Ss were divided according to the peak scores they obtained on the MMPI. Fifteen Ss with a peak score on *Pd* and five Ss with a peak score on *Ma* followed by a secondary peak on *Pd* formed the *PPd* group (Peak Psychopathic Deviate).³ The 28 remaining Ss formed the *NPPd* group (Nonpeak Psychopathic Deviate). The major peak scores of Ss comprising the *NPPd* group were *Hy* (*N* = 8), *Sc* (*N* = 7), and *Ma* (*N* = 6). Nine of the *PPd* Ss and 12 of the *NPPd* Ss had at least one *T* score of 70 or above. Table 1 presents means and standard deviations for the two groups on three matching variables and indicates that the groups did not differ in age, vocabulary level (measured by the 40-item Shipley-Hartford Scale), or level of manifest anxiety (measured by the 50-item Taylor *MA* scale, Scale 16 in Appendix I, Dahlstrom & Welsh, 1960).⁴

Cartoon Stimuli

Three clinical psychologists rated independently the content of hundreds of cartoons. The variety of cartoon sources ranged from the *New Yorker Twenty-Fifth Anniversary Album* (1951) to *Best Cartoons of the Year 1965* (Lariar, 1965) to *College Laughs* (1964). From the cartoons which were categorized with perfect agreement, *E* selected 40 cartoons with hostile content, 40 cartoons with sexual content,

³ Following customary procedure, Scales *Mf* and *Si* were excluded from consideration as peak scores. The *PPd* group contained Ss with two different profiles because, in a normal population, an *MaPd* sequence is commonly thought to be essentially the same as a *PdMa* sequence (Dahlstrom & Welsh, 1960). Examination of *PPd* profiles indicates that the group as a whole is most accurately described as consisting of normal individuals with psychopathic tendencies. A majority (11) of the Ss in *PPd* had a *PdMa* or *MaPd* profile. Normal persons with these profile patterns are expected to "show clear manifestations of psychopathic behavior [Dahlstrom & Welsh, 1960, p. 192]." Considering only the first digit of the profile, "peak scores on Scale four (*Pd*) almost without regard to elevation of the profile [provide] evidence of lack of social conformity or self-control and a persistent tendency to get into scrapes [Dahlstrom & Welsh, 1960, p. 188]."

⁴ It was important for anxiety level (Doris & Fierman, 1956) and vocabulary level (Holmes, 1968) to be controlled because of previously demonstrated relationships with humor of the type used in this study.

and 30 cartoons with nonsense content. In addition, *E* inserted 10 cartoonlike stimuli which were out of context and not intended to be funny when in isolation. These stimuli were introduced in order to keep *S* honest in the Immediate Reaction situation. The cartoons were split randomly into two equivalent forms, A and B, with some attempt made to separate cartoons with highly similar specific content. Form A and Form B each had 20 cartoons with hostile and 20 with sexual content, 15 cartoons with nonsense content, and 5 cartoonlike stimuli that were not funny.

Categorization Measures

Four by six photographic reproductions of the stimuli were printed and each was encased in transparent plastic. Following precedent (Byrne, 1956), *E* requested Ss to sort the cartoons into four piles of 15 each, ranging from Most Funny, to Next, to Next, to Least Funny. Order of the cartoons was randomized for each *S*. Initial measures were obtained by summing the scores for each content area, with Most Funny = 4, Next = 3, Next = 2, and Least Funny = 1.

Immediate Reaction Measures

A standard 35-mm. slide of each stimulus was obtained for use in a Kodak Carousel projector. The *S* was seated in a dark room in front of a projection screen and was asked to rest his right arm on a table of appropriate height. A small box containing one large protruding butterfly-shaped knob and one push button was placed on this table. After *S* was told to push the button in order to have a slide projected through a small window, *E* left the room. A series of 12 slides, including some practice cartoons similar to those contained in Forms A and B, introduced *S* to the mechanics of turning the knob on a continuum of 180 degrees in order to register his appreciation of the cartoons. The *S* had to return the butterfly knob to the zero level before pushing the button for the next slide. Following this, *S* was instructed to turn the knob an appropriate distance as soon as he got the point of a cartoon, then switch to the next cartoon as soon as he was finished with

the previous one. Since some stimuli were not funny at all, he was told simply to switch to the next cartoon without turning the knob whenever he thought a stimulus did not contain any humor. After completion of the 12 training slides, either Form A or Form B was presented, with *S* totally controlling the rate of cartoon presentation. Order of the cartoons was random, with the restriction that three hostility, three sex, and three nonsense or not-funny stimuli occur in each consecutive set of nine. Once order had been determined, it remained constant throughout the study.

A continuous recording of *S*'s responses was obtained in the slide projector room by means of a Heath Model EU-W 20 strip chart recorder moving at 12 inches per minute. Movement of the push button, which activated the projector only when the butterfly knob was at zero level, was indicated on the chart paper by movement of the recorder pen in a negative direction.

Two measures were obtained in the Immediate Reaction situation. *Latency* of response refers to the time between presentation of a cartoon and reaction to it. *Amplitude* of response refers to the distance the knob was turned by *S*, with the two poles of the continuum being zero for not funny and 80 for very funny.

Responses to the not-funny stimuli are not included in the Results section. For the most part, these stimuli were seen as not funny and thereby served the purpose of compelling *S* to respond to each cartoon as a separate stimulus.

Procedure

The *Ss* were scheduled in groups of two to six each, tested individually, and then scheduled for a second session to be conducted about 1 week later. Each of the techniques, Immediate Reaction and Categorization, plus each of the forms, A and B, were exposed to *Ss* in an alternate manner: one technique and one form on the first day, the other technique and form on the second day. Several tests besides the MMPI and Shipley-Hartford Vocabulary List were administered. Each session lasted approximately 1½ hours.

Transformations of Data

For each *S*, mean responses to sex, hostility, and nonsense cartoons were summed and the proportion of that sum contributed by sex, hostility, and nonsense means was calculated. The proportions, which underlie the analyses of amplitude and categorization responses reported in Table 3, represent "ipsative" rather than "normative" measures (Cattell, 1944). Although the distinction between "ipsative" and "normative" measures is important to personality theorists and diagnosticians interested in humor (Tollefson & Cattell, 1963-1964), it is largely irrelevant in the present study because overall liking for the cartoons used appears to be the same for *PPd* and *NPPd* *Ss* (Table 2, normative data).

RESULTS

Table 2 presents a summary of latency responses to the 55 cartoons containing humor. Nonsense cartoons were responded to more quickly than other cartoons, perhaps because there were relatively fewer cartoons that required reading in the nonsense category. All cartoons were responded to more quickly by *PPd* than by *NPPd* *Ss*. On the average, *PPd* *Ss* responded more than 1½ seconds faster than *NPPd* *Ss*. This relatively quick response time is not attributable to hypomania on the part of the peak-*Ma* *Ss* included in *PPd* because the mean of the five peak-*Ma* *Ss* in this group was higher than the mean of the remaining 15 *PPd* *Ss*.

Although measures of cartoon comprehension were not obtained, it seems likely for two reasons that *PPd* *Ss* actually saw a humorous point—not necessarily *the only* point—to the cartoons more rapidly than *NPPd* *Ss*. First, all *Ss* reported turning the knob as soon as the point to a cartoon was seen, and second, the obtained difference of more than 1½ seconds is far greater than would be expected on the basis of reaction-time differences in initiating a psychomotor act (turning the knob) in response to a signal. It has not been demonstrated, however, that the obtained differences in latency of response reflect comprehension of specific points in the cartoons or that the obtained differences are specific to humorous stimuli. They may relate instead to more general response tendencies, to stimulus characteristics not uniquely humorous, or to aspects of the measurement situation per se that differentially influence *PPd* and *NPPd* *Ss*.

Table 2 also summarizes amplitude responses obtained in the Immediate Reaction situation. The nonsignificant main effect indicates that there is no evidence to suggest that *PPd* and *NPPd* *Ss* differ in the overall extent to which they appreciate humor. The significant interaction effect, however, indicates that *PPd* and *NPPd* *Ss* differ in their appreciation of particular types of humor. Examination of the mean scores reveals that *PPd* *Ss* appreciated sex and hostility cartoons more than nonsense cartoons, with the difference between sex and nonsense cartoons being sta-

TABLE 2
IMMEDIATE REACTION RESPONSES TO HUMOR

Source	df	Amplitude		Latency	
		MS	F	MS	F
Between Ss	1	193.69	.43	95.19	8.05***
PPd, NPPd	46	446.90		11.82	
Ss within groups					
Within Ss					
Sex, hostility, non-sense cartoons	2	13.97	.20	9.83	15.60****
PPd, NPPd X Sex,					
Hostility, Non-sense Cartoons	2	250.61	3.64**	1.54	2.44*
Residual	92	68.89		.63	
Mean amplitude					
		Sex	Hostility	Nonsense	
PPd (N = 20)		34.85 _b	33.95 _{ab}	28.90 _a	
NPPd (N = 28)		28.93 _a	29.93 _{ab}	31.79 _{ab}	
Mean latency					
		Sex	Hostility	Nonsense	
PPd (N = 20)		6.42 _{ab}	6.63 _b	6.07 _a	
NPPd (N = 28)		8.40 _d	8.32 _d	7.34 _c	

Note.—For amplitude and latency considered separately, with $S_m = \sqrt{\text{Residual}/20}$, cells containing the same subscripts are not significantly different from each other at the .05 level by Duncan multiple-range test (Duncan, 1955).

* $p < .10$.
** $p < .05$.
*** $p < .01$.
**** $p < .001$.

tistically significant by range test. *NPPd* Ss appear to have a more balanced appreciation of the three cartoon types.

Mean scores in Table 3 highlight the pattern of differences revealed by the significant interaction effect in Table 2.⁵ (Sex and hostility scores were combined in Table 3 because inspection of means and standard deviations revealed only trivial differences between them. For sex and hostility, respectively, mean amplitude proportions were .354 and .350 for *PPd*, and .320 and .334 for *NPPd*; mean categorization proportions were .349 and .346 for *PPd*, and .333 and .342 for *NPPd*.) In both Immediate Reaction and Categorization situations, *PPd* Ss preferred sex plus hostility cartoons significantly more than nonsense cartoons. No particular preference was exhibited by *NPPd* Ss, with the direction of statistically insignificant differences being reversed in the

⁵ Note that main effect differences between *PPd* and *NPPd* in Table 3 are artifacts caused by dividing sex plus hostility by two in the presence of differences in *PPd*, *NPPd* preferences for sex, hostility, and nonsense cartoons. Interpretations of Table 3 in the text have taken this into account.

two measurement situations. The pattern in the data, then, is that the preference for sex plus hostility cartoons shown by *PPd* Ss is accompanied, in the absence of a preference by *NPPd* Ss, by differences between *PPd* and *NPPd* Ss in cartoon preference. Statistically significant differences are that *PPd* Ss enjoyed nonsense cartoons less than *NPPd* Ss; as indicated in the amplitude proportion data (Table 3), and enjoyed sex cartoons more, as indicated in Table 2.

Although, as a group, *PPd* Ss appreciated sex and hostility cartoons to an equal extent, individual preferences in the *PPd* group are related to variables measured by the *Mf* (masculinity-femininity) scale of the MMPI. For the *PPd* group, the Pearson product-moment correlations between the *Mf* scale of the MMPI and preferences for sex cartoons are $-.47$ for Immediate Reaction (ipsative data) and $-.40$ for Categorization. A correlation of $-.44$ is significantly different from a zero correlation at the .05 level of confi-

TABLE 3
CATEGORIZATION AND AMPLITUDE (PROPORTIONS)
RESPONSES TO HUMOR

Source	df	Categorization		Amplitude (Proportions)	
		MS	F	MS	F
Between Ss <i>PPd</i> , <i>NPPd</i>	1	679.50	2.44	3,565.04	6.43*
Ss within groups	46	279.04		554.59	
Within Ss					
Sex plus hostility, nonsense cartoons	1	15,478.76	6.53**	4,017.10	.96
<i>PPd</i> , <i>NPPd</i> X Sex plus Hostility, Nonsense Cartoons	1	5,743.58	2.42	32,681.38	7.80**
Residual	46	2,370.56		4,191.35	
Mean categorization proportions					
		Sex plus hostility		Nonsense	
<i>PPd</i> (N = 20)		.348 _b		.304 _a	
<i>NPPd</i> (N = 28)		.338 _b		.325 _{ab}	
Mean amplitude proportions					
		Sex plus hostility		Nonsense	
<i>PPd</i> (N = 20)		.352 _b		.296 _a	
<i>NPPd</i> (N = 28)		.327 _{ab}		.346 _b	

Note.—For categorization and amplitude considered separately, with $S_m = \sqrt{\text{Residual}/20}$, cells containing the same subscripts are not significantly different from each other at the .05 level by Duncan multiple-range test (Duncan, 1955).

* $p < .05$.
** $p < .01$.

dence. The interpretation is that *PPd* Ss who have more masculine interests prefer sex cartoons more than hostility and nonsense cartoons. With *NPPd* Ss this does not appear to be true. The comparable correlations for the *NNPd* Ss are $-.12$ for Immediate Reaction and $-.16$ for Categorization. These small, nonsignificant correlations suggest that the relationship between masculine interests and appreciation of sexual cartoons is largely or totally confined to the *PPd* Ss.

DISCUSSION

Since it is commonly assumed that hostile and sexual impulses are expressed in hostile and sexual humor, the cartoon stimuli in this study were initially categorized into sex, hostility, and nonsense. The five major substantive findings are the following: (a) *PPd* and *NPPd* Ss, overall, appreciated humor to an equal extent; (b) *PPd* Ss preferred sexual and hostile humor more than nonsense humor; (c) *PPd* Ss preferred sexual (and hostile) humor more and nonsense humor less than *NPPd* Ss; (d) *PPd* Ss with strong masculine interests preferred sexual cartoons more than *PPd* Ss with less strong masculine interests; (e) *PPd* Ss appeared to see the point to all cartoons more quickly than did *NPPd* Ss. These findings with *PPd* Ss fit well with straightforward expectations of behavior derived from knowledge of psychopathic persons who are not experiencing situational anxiety. Psychopathic persons are generally described as shallow, glib individuals who lack inhibition and seek immediate gratification of impulses (Coleman, 1956). It is not surprising, therefore, to discover that Ss with a psychopathic tendency appear to respond quickly or glibly to humor, prefer humor which involves impulse gratification, and uninhibitedly prefer heterosexually relevant humor more if they are actually more interested in the opposite sex. The results should be generalized cautiously, however, because psychiatric technicians have been infrequently studied and it may be that their population characteristics interact somehow with the variables measured in the present study.

Other investigators who have used similar stimulus materials have also discovered positive correlations between strength or expres-

sion of impulses in a humorous context and strength or expression of such impulses in a nonhumorous context. The variety of situations studied ranges from observation of actual aggressive behavior (Byrne, 1956) to aggressive sentiments expressed on paper-and-pencil tests (Hetherington & Wray, 1964; Murray, 1934) to content of TAT stories (Grziwok & Scodel, 1956) to experimentally aroused hostility (Dworkin & Efran, 1967; Strickland, 1959). It is also becoming evident that high inhibition concerning aggression interferes with apparent enjoyment of hostile humor. Gollob and Levine (1967) and Singer, Gollob, and Levine (1966) demonstrated that experimental arousal of aggressive inhibition decreases preference for aggressive humor. Hetherington and Wray (1964) demonstrated that dissolving inhibition in liquor increases preferences for hostile humor in high-aggression, high-inhibition Ss.

The results of the present study, in which high-impulse, low-inhibition (*PPd*) Ss responded more appreciatively to impulse-related humor, are consistent with what would be expected from the studies mentioned above. Furthermore, obtained differences between *PPd* and *NPPd* Ss suggest that it may be necessary to distinguish between these two kinds of Ss when attempting to understand the dynamics of humor. Perhaps, for *PPd* Ss, responding to humor is simply one of many alternative ways to act out impulses. In the absence of deep-rooted conflicts concerning acting out, there would be little reason for humor to serve as an outlet for bottled-up emotions. With *NPPd* Ss, for whom acting out in general is more difficult, responding to humor may be one of the few available avenues of expression. If so, complex and subtle patterns of psychodynamic determination may pertain to *NPPd* Ss, but not, under ordinary circumstances, to *PPd* Ss.

The attempt to measure humor-being-perceived worked successfully in the present study. The Ss readily accepted the Immediate Reaction situation. With only one exception, in response to an open-ended question, Ss first commented on the cartoons or on their own sense of humor before commenting on the response task. They appeared to pay attention to the stimuli rather than to the appa-

ratus or the measuring technique. Hence, it would seem that neither elaborate warm-up procedures nor "special" laboratory environments are necessary in order to obtain reasonable measures of humor-being-perceived. The rapport usually established to obtain an S's cooperation apparently is sufficient.

Measures of humor-being-perceived appear to possess a substantial potential for use in laboratory research. The type of measurement situation used in Immediate Reaction lends itself well to systematic laboratory study by means of tightly controlled manipulation of stimulus, environmental, and S variables. For example, stimuli might be varied in their complexity, content, and subtlety; social environment might be varied by introducing the illusory presence of other Ss through tape-recordings; emotional states might be manipulated by visual or auditory presentation of nonhumorous material.

Measures of humor-being-perceived also appear to possess potential value for application. As a measure of emotional functioning, latency might reveal perceptual sensitization or blocking in areas of conflict or vulnerability. As a measure of intellectual functioning, latency might reveal speed of cognitive processing as a function of stimulus complexity. In both applications, with Ss kept honest by presentation of not-funny cartoonlike stimuli, amplitude measures would provide an indication of whether or not S actually saw the point. Amplitude measures also, of course, would indicate degree of preference for different stimuli. If knowledge of the contribution of emotional and intellectual factors were available, then instruments might be developed which would be relatively free from the effects of test anxiety, test-taking styles, and test-wise Ss.

REFERENCES

- ALLPORT, G. W. *The nature of prejudice*. Boston: Addison-Wesley, 1954.
- BYRNE, D. The relationship between humor and the expression of hostility. *Journal of Abnormal and Social Psychology*, 1956, 53, 84-89.
- CATTELL, R. B. Psychological measurement: Ipsative, normative, and interactive. *Psychological Review*, 1944, 51, 292-303.
- COLEMAN, J. D. *Abnormal psychology and modern life*. Chicago: Scott, Foresman, 1956.
- College Laughs, 1, No. 36 (August 1964). Holyoke, Mass.: Major Magazines, 1964.
- DAHLSTROM, W. G., & WELSH, G. S. *An MMPI handbook*. Minneapolis: University of Minnesota Press, 1960.
- DORIS, J., & FIERMAN, E. Humor and anxiety. *Journal of Abnormal and Social Psychology*, 1956, 53, 59-62.
- DUNCAN, D. B. Multiple range and multiple F tests. *Biometrics*, 1955, 11, 1-42.
- DWORKIN, E. S., & EFRAN, J. S. The angered: Their susceptibility to varieties of humor. *Journal of Personality and Social Psychology*, 1967, 6, 233-236.
- FLUGEL, J. C. Humor and laughter. In G. Lindzey (Ed.), *Handbook of social psychology*. Cambridge, Mass.: Addison-Wesley, 1954.
- GOLLOB, H. F., & LEVINE, J. Distraction as a factor in the enjoyment of aggressive humor. *Journal of Personality and Social Psychology*, 1967, 5, 368-372.
- GRZIWOK, R., & SCODEL, A. Some psychological correlates of human preferences. *Journal of Consulting Psychology*, 1956, 20, 42.
- HETHERINGTON, E. M., & WRAY, N. P. *Journal of Abnormal and Social Psychology*, 1964, 68, 685-689.
- HOLMES, D. S. Sensing humor: Latency and amplitude of response related to vocabulary level. Unpublished manuscript, 1968.
- JENNESS, A. Personality dynamics. In P. R. Farnsworth, O. McNemar, & Q. McNemar (Eds.), *Annual Review of Psychology*, 1964, 13, 479-514.
- KAMBOUROPOVLOV, P. Individual differences in the sense of humor. *American Journal of Psychology*, 1926, 37, 288-297.
- LARIAR, L. (Ed.) *Best cartoons of the year 1965*. New York: Crown, 1965.
- LUBORSKY, L. B., & CATTELL, R. B. The validation of personality factors in humor. *Journal of Personality*, 1947, 15, 283-291.
- MURRAY, H. A., JR. The psychology of humor. *Journal of Abnormal and Social Psychology*, 1934, 25, 66-81.
- New Yorker twenty-fifth anniversary album*. New York: Harper, 1951.
- SINGER, D. L., GOLLOB, H. R., & LEVINE, J. Inhibitions and the enjoyment of aggressive humor: An experimental investigation. Paper presented at the meeting of the Eastern Psychological Association, New York, April 1966.
- STRICKLAND, J. F. The effect of motivation arousal on humor preferences. *Journal of Abnormal and Social Psychology*, 1959, 59, 278-281.
- TOLLEFSON, D. L., & CATTELL, R. B. *Handbook for the IPAT Humor Test of Personality*. Champaign, Ill.: The Institute for Personality and Ability Testing, 1963-1964.

(Received June 8, 1968)

PSYCHIATRIC HOSPITAL EXPERIENCE AND ATTITUDES TOWARD "MENTAL ILLNESS"

JEAN J. SMITH¹

Kings View Hospital, Reedley, California

Prior to and following psychiatric hospital experience, an experimental group composed of patients, student nurses, and hospital employees was tested with a five-factor attitude scale, *Opinions about Mental Illness*. The group showed significant decreases in attitudes of authoritarianism and social restrictiveness toward mental patients. Attitude changes appeared not to be related to differences in indoctrination or time spent in the hospital. All significant attitude changes in the experimental group were in the direction of greater conformity to the attitudes of professional treatment staff or greater favorableness. A control group showed a significant decrease in belief in interpersonal etiology, but otherwise showed no significant changes.

Earlier studies have shown that during psychiatric hospital experience, attitudes of student nurses toward the "mentally ill" change in a favorable direction (Hicks & Spaner, 1962) and those of patients change toward conformity to the attitudes of the treatment staff (Manis, Houts, & Blake, 1963). Since Struening and Cohen (1963) have analyzed attitudes toward "mental illness" into five factors, it is of interest to determine whether changes in attitudes are uniform on all five factors or are selective. The first factors are described as follows by Cohen and Struening (1962):

Factor A—Authoritarianism. . . . a gestalt made up of authoritarian submission and anti-intracception with a view of the mentally ill as a class inferior to normals and requiring coercive handling.

Factor B—Benevolence. . . . a kindly, paternalistic view towards patients whose origin is in religion and humanism rather than a scientific or professional dogma. It is encouraging and nurturant, but still acknowledges some fear of patients.

Factor C—Mental Hygiene Ideology. . . . the idea that mental patients are much like normal people, differing from them perhaps in degree, but not in kind, in sharp contrast with the Factor A orientation.

Factor D—Social Restrictiveness. . . . the belief that mental illness is a threat to society which must be met by some restriction in social functioning both during and following hospitalization.

Factor E—Interpersonal Etiology. . . . a belief that mental illness arises from interpersonal experience, particularly deprivation of parental love and attention during childhood . . . or more generally the mental health of those *in loco parentis*. . . Some-

what less central is a belief that abnormal behavior is motivated; e.g., mental illness is an avoidance of problems . . . [pp. 352-355].

Lewis and Cleveland (1966) indicate that student nurses, following an 8-week psychiatric experience, exhibited attitudinal changes on Cohen and Struening's scales. They found a significant decrease in Authoritarianism, and increases in Mental Hygiene Ideology and Interpersonal Etiology. No consistent changes were observed for Benevolence or Social Restrictiveness.

The purpose of the present study was to investigate further the specific nature of change in the attitudes of several groups after varying lengths of exposure to psychiatric hospital experience.

METHOD

Subjects

Three experimental subgroups and two control subgroups were included, totaling 129 experimental and 103 control Ss. One experimental subgroup consisted of 65 patients consecutively admitted, for the first time, to a small, church-sponsored, private psychiatric hospital whose therapeutic community treatment approach maximizes interpersonal contact at all levels. This patient subgroup was heterogeneous with respect to diagnosis, education, cultural background, and religious orientation. The median age of the group was 38 years, and the range was 16-67 years. Length of hospitalization varied from 3 to 34 weeks, with a median of 10 weeks. Excluded as Ss were patients who had previous admissions, who were too disturbed at the time of admission to cooperate in giving data, and those from whom it was impossible to obtain posttest data, usually because of abrupt departure. These constituted 30% of the total sample.

¹Requests for reprints should be sent to J. J. Smith, Kings View Hospital, P. O. Box 631, Reedley, California 93654.

A second experimental subgroup included 33 student nurses who underwent a psychiatric affiliation in the same hospital. All student nurses spent the same number of days in the hospital, although for some the experience was spread over a 16-week period, while for others it was 8 weeks. Their ages ranged from 20 to 40 years, with a median of 21.

The third experimental subgroup consisted of 31 conscientious objectors and Mennonite Voluntary Service workers (VS) employed as ward personnel by the hospital. These employees spent 1-2 years at the hospital, depending upon the nature of their assignments under Selective Service or Voluntary Service. Their age range was 19-35 years; the median was 22. The median age of the total experimental group was 23 years with a range from 16 to 67.

The control group included 81 students enrolled in an introductory psychology course at the same college as that attended by the student nurses and 22 Mennonite Voluntary Service personnel (VS controls) who were assigned to nonpsychiatric service in other locations. Ages of the students ranged from 19 to 58 years with a median of 20, while the range of the VS controls was 20-48 years, with a median of 23. The median age of the total control group was 21 years with a range from 19 to 58.

Hypotheses and Procedure

A questionnaire consisting of the items from Cohen and Struening's (1962) Opinion about Mental Illness (OMI) was administered on a test-retest basis to all members of the experimental and control groups. The first hypothesis to be tested was that the pretest-posttest scores of the experimental group would reflect the findings of Manis, Houts, and Blake (1963), exhibiting shifts in the direction of conformity to the attitudes of the hospital treatment staff. The second hypothesis was that such convergence, should it occur, would also support the findings of Hicks and Spaner (1962) that the attitudes of student nurses toward psychiatric patients shifted in a favorable direction.

In order to provide a base line for the evaluation of attitudinal shifts, as well as to provide a criterion of "favorableness," the OMI was administered to the clinical treatment staff of the hospital (three psychologists, five psychiatrists, and four social workers). The mean scores of this group on the five OMI factors were then used as a base line against which to evaluate pretest and posttest scores of both the experimental and control groups for degree of shift and for relative favorableness or unfavorableness.

Pretest scores were obtained from all the members of the experimental group upon their arrival at the hospital. Patients were tested immediately after admission, student nurses during the first day of affiliation, and employees when they started work. Posttest scores were obtained from the patient and student nurse subgroups just prior to their departure from the hospital, either at the time of discharge or upon termination of their hospital affiliation.

The members of the employee subgroup, by reason

of their status as volunteers or conscientious objectors, were committed to work at the hospital for predetermined periods of from 1 to 2 years. Consequently, it was decided to retest them at varying time intervals in order to test the second hypothesis that degree of change in attitudes would be positively related to time spent in the hospital. Four retest intervals were employed. One-fourth were retested after they had been in the hospital for 1 month; one-fourth were retested after 6 months; one-fourth after 12 months; and the remaining fourth after 24 months. Assignment of individuals to each of the four retest intervals was made on a rotational basis, that is, S_1 after one month, S_2 after 6 months, S_3 after 12 months, etc. Obviously, however, only S_s who were known to be assigned to the hospital for 24-month tours could be assigned to the fourth interval, and adjustments in the distribution of S_s among the four retest intervals were made accordingly.

The control S_s were tested in groups. The student control S_s received the OMI pretest at the beginning of their introductory course in psychology, and the posttest 14 weeks later before beginning the mental health unit of their course. The VS control S_s completed the pretest as part of their orientation for nonpsychiatric work assignments. Posttest scores for these S_s were obtained by canvassing them by mail 1 year later. Since at this time they were widely scattered throughout the world, posttest scores were obtainable for only approximately 60% of the original VS control subgroup. This selectivity of response may have biased the subgroup's results, possibly accounting for its change on one attitude dimension, as seen below.

A third hypothesis predicted that the varying degrees and kinds of indoctrination to which the experimental subgroups were subjected would be reflected by varying degrees of convergence of posttest means toward the mean of the treatment staff on each of the five factors. Since the student nurses were subjected both to hospital exposure and to systematic instruction, they were expected to show the most change. Employees were subjected to in-service training efforts which, while not as systematic or as intensive as those of the student nurses, nevertheless suggested that employees might change more than patients. Hence, the predicted order for degree of change was student nurses > employees > patients. Since the control subgroups were subjected neither to indoctrination nor to exposure to psychiatric patients, their posttest means were not expected to converge toward the treatment staff means or to change in degree of favorableness to patients.

All questionnaires were scored separately for each of Cohen and Struening's five factors. Scoring was based on the published factor loadings of 53 items (Cohen & Struening, 1962); items were scored in the direction of their positive or negative factor loadings for the respective scales. One item, "The patients of a mental hospital should have something to say about the way the hospital is run," received a high-negative loading for Factor A and a high-

positive loading for Factor C. This item was scored only for Factor C. A 6-point Likert forced-choice response continuum was used in the questionnaire. Response alternatives were: Agree very much, Agree on the whole, Agree a little, Disagree a little, Disagree on the whole, and Disagree very much. Responses were scored from 0 (extreme disagreement) to 7 (extreme agreement), omissions being scored 4. Item scores were summed algebraically with negative values assigned to items having negative factor loadings.

RESULTS AND DISCUSSION

Pretest Group Differences

Inspection of Table 1 reveals that the mean pretest OMI score of the experimental group was significantly higher ($p < .01$)² than that

²Significance of all pretest intragroup differences was evaluated by means of two-tailed *t* tests, for independent means, since no predictions had been made about their direction. Pretest-posttest differences were evaluated for significance by means of one-tailed tests for correlated means since their directions had been predicted.

of the control group on Mental Hygiene Ideology and significantly lower ($p < .01$) on Social Restrictiveness. The control group was higher on Authoritarianism, but the difference was of borderline significance ($.05 > p > .01$).

The mean pretest scores of all subgroups were higher than those of the treatment staff for Authoritarianism, Benevolence, and Social Restrictiveness, and below on Mental Hygiene Ideology. These differences were all significant at or beyond the .01 level. The student nurses were significantly less authoritarian on pretest ($p < .001$) than were the other experimental or control subgroups, but there were no consistent pretest differences between subgroups on the remaining four factors.

Pretest-Posttest Group Changes

Comparison of pretest and posttest OMI means and standard deviations is also seen in

TABLE 1
PRETEST AND POSTTEST OMI MEANS AND STANDARD DEVIATIONS OF ALL GROUPS

Group	Authoritarianism		Benevolence		Mental hygiene ideology		Social restrictiveness		Interpersonal etiology		N
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test	
Treatment staff											12
<i>M</i>	21.2	—	40.1	—	51.0	—	7.8	—	24.9	—	
<i>SD</i>	8.5		3.2		4.7		3.6		6.2		
Experimental group											129
<i>M</i>	47.8**	40.1	44.3	43.5	44.7	45.2	14.2**	11.8	23.2	24.1	
<i>SD</i>	21.6	20.2	6.3	6.4	7.8	7.9	8.7	7.6	7.1	7.1	
Nurses											33
<i>M</i>	31.1*	26.4	44.4	43.4	45.0	46.5	12.1*	10.1	22.0	24.6	
<i>SD</i>	14.2	11.8	5.6	5.4	5.6	5.4	5.7	5.8	5.3	5.4	
Patients											65
<i>M</i>	54.1**	44.5	43.5	43.3	45.0	45.4	15.1*	12.9	22.7	23.0	
<i>SD</i>	22.7	21.4	6.8	7.6	8.7	8.8	10.6	8.3	7.5	8.1	
Employees											31
<i>M</i>	52.4**	45.8	45.9*	44.0	43.9	43.4	14.6**	11.1	25.6	25.8	
<i>SD</i>	15.5	17.9	5.2	4.8	7.6	8.0	6.3	7.3	7.4	5.7	
Control group											103
<i>M</i>	54.0*	51.2	44.4	44.0	40.5	40.2	20.8*	19.4	24.1**	22.2	
<i>SD</i>	17.0	16.6	5.6	5.1	7.5	7.9	8.0	7.4	6.6	6.5	
Students											81
<i>M</i>	53.5**	50.1	44.3	44.3	39.8	39.1	20.9	19.7	23.0*	21.5	
<i>SD</i>	16.3	15.1	5.9	5.0	7.8	7.8	7.8	7.2	6.4	6.2	
Voluntary service											22
<i>M</i>	56.1	55.4	44.4	42.8	43.2	44.0	20.3	18.4	28.1**	24.8	
<i>SD</i>	19.5	20.6	4.5	5.0	5.8	7.2	8.7	8.3	5.7	5.8	

* $p = .05$, one-tailed *t* test.
** $p = .01$, one-tailed *t* test.

Table 1. Overall changes from pretest to posttest were clearly selective for the various factors. The total experimental group showed significant decreases ($p < .01$) on Authoritarianism and Social Restrictiveness, converging on the treatment staff norms in agreement with the findings of Manis, Houts, and Blake (1963). The total group showed borderline decreases ($p < .05$) on Authoritarianism and Social Restrictiveness, and a significant ($p < .01$) decrease on Interpersonal Etiology. The experimental group remained significantly higher on Mental Hygiene Ideology and lower on Social Restrictiveness than the control group ($p < .01$). In addition, at posttest the control group was significantly more authoritarian than the experimental group, and the difference on Interpersonal Etiology approached significance ($.05 > p > .01$).

Most of the posttest means of the experimental and control subgroups remained significantly above those of the treatment staff on Authoritarianism, Benevolence, and Social Restrictiveness, and significantly below on Mental Hygiene Ideology. Exceptions were in the student nurse, patient, and VS subgroups. The differences between the student nurses and the treatment staff on Authoritarianism, Benevolence, and Social Restrictiveness were no longer significant. In the patient and VS subgroups the posttest means on Benevolence were no longer significantly above that of the treatment staff.

Individual subgroups showed varying patterns of change. The student nurse subgroup showed a borderline decrease in Authoritarianism and Social Restrictiveness, and a borderline increase in Interpersonal Etiology. This partially supported the findings of Lewis and Cleveland (1966), who found a decrease in Authoritarianism and increases in Mental Hygiene and Interpersonal Etiology, among student nurses. The patient subgroup showed a significant decrease in Authoritarianism, and a borderline decrease in Social Restrictiveness. The employee subgroup decreased to a significant degree in Authoritarianism and Social Restrictiveness, and to a borderline degree in Benevolence.

Three changes in the control subgroups are of interest. These are the significant decrease

in Authoritarianism and the borderline decrease in Interpersonal Etiology shown by the student controls, and the significant decrease in Interpersonal Etiology shown by the VS controls. The first two may reflect the influence of the introductory psychology course to which the student controls were exposed between pretest and posttest. However, Costin and Kerr (1962) found that exposure to an abnormal psychology class resulted in student OMI changes in the direction of lower Authoritarianism, Benevolence, and Social Restrictiveness, and increased Interpersonal Etiology. The significant decrease in Interpersonal Etiology in the VS controls may reflect their experience with foreign customs and living conditions, and increased awareness of the importance of social and economic factors in the etiology of emotional disturbance. However, evidence for this interpretation is lacking. Further study might reveal whether or not the changes observed in these control subgroups are fortuitous.

Order of Change

The data yielded no evidence that degree of change followed the predicted order of student nurses > employees > patients. A chi-square for the relative ranks of the three subgroups on all five factors failed to approach significance ($p > .05$). The borderline increase in Interpersonal Etiology exhibited by the student nurses may reflect the influence of the formal instruction which supplemented their hospital experience, although the data neither support nor disprove this speculation. Indeed, the apparent predominance of attitudinal change (Authoritarianism, Benevolence, Social Restrictiveness) over ideological change (Mental Hygiene Ideology, Interpersonal Etiology) within the experimental group suggests that the changes which did occur may be attributable to the effects of increased personal familiarity with patient behavior, rather than to those of indoctrination or instruction.

Length of Hospital Experience

In order to test the hypothesis of a relationship between degree of change and length of experience in the hospital, an analysis of

variance was carried out on the scores of the members of the employee subgroup, who had been retested after 1, 6, 12, or 24 months. Since this subgroup exhibited significant change only on Authoritarianism and Social Restrictiveness, only the scores of these factors were analyzed. The F values obtained were 1.30 and .80 for the two factors, respectively, lending no support to the hypothesis since they were far below the values required for statistical significance. These findings suggest that attitude changes may well have occurred in the early stages of Ss' hospital experience and that they were not cumulative with time.

However, since the small N (31) calls into question the power of the F test of significance in this instance, this conclusion must be tentative. Nevertheless, the present result is of interest in view of recent studies such as that of Muench, Tutko, and Stuckey (1967), and that of Keith-Spiegel and Spiegel (1967), in which attitudes of students toward psychiatric patients have been shown to change as result of rather extensive exposure to a hospital environment. The present results suggest that change might be induced by considerably shorter periods of exposure than the approximately semester-long periods employed in these studies. Further study with more Ss is needed.

REFERENCES

- COHEN, J., & STRUENING, E. L. Opinions about mental illness in the personnel of two large mental hospitals. *Journal of Abnormal and Social Psychology*, 1962, 64, 349-360.
- COSTIN, F., & KERR, W. D. The effects of an abnormal psychology course on students' attitudes toward mental illness. *Journal of Educational Psychology*, 1962, 53, 214-218.
- HICKS, J. M., & SPANER, F. E. Attitude change and mental hospital experience. *Journal of Abnormal and Social Psychology*, 1962, 65, 112-120.
- KEITH-SPIEGEL, P., & SPIEGEL, D. E. Changes in attitude towards mental illness among volunteer students following experience in a neuropsychiatric hospital. Paper presented at the meeting of the California State Psychological Association, Coronado, January 1967.
- LEWIS, D. L., & CLEVELAND, S. E. Nursing students' attitudinal changes following a psychiatric affiliation. *Journal of Psychiatric Nursing*, 1966, 4, 223-231.
- MANIS, M., HOUTS, P. S., & BLAKE, J. B. Beliefs about mental illness as a function of psychiatric status and psychiatric hospitalization. *Journal of Abnormal and Social Psychology*, 1963, 67, 226-233.
- MUENCH, G., TUTKO, T., & STUCKEY, R. The impact upon students of a ten-week mental hospital visitation program. Paper presented at the meeting of the Western Psychological Association, San Francisco, May 1967.
- STRUENING, E. L., & COHEN, J. Factorial invariance and other psychometric characteristics of five opinions about mental-illness factors. *Educational and Psychological Measurement*, 1963, 23, 289-298.

(Received June 13, 1968)

DEVELOPMENT OF AN EMPATHY SCALE¹

ROBERT HOGAN²

Johns Hopkins University

The concept of empathy—the intellectual or imaginative apprehension of another's condition or state of mind—is central for understanding a broad range of social phenomena including, in particular, moral development. Within this latter context, an empathic disposition can be regarded as the capacity to adopt a broad moral perspective, that is, to take "the moral point of view." This paper discusses the development of a 64-item self-report measure of empathy, constructed by comparing the responses of groups with high- and low-rated empathy, using the combined MMPI-CPI item pool. After providing evidence concerning the scale's reliability and validity, an attempt is made to show its relevance for specifically moral conduct by relating empathy scale scores to real life indexes of socially appropriate behavior and to certain previously well-validated measures of personality.

Some form of empathic disposition, role-taking ability, or social sensitivity is assumed by all approaches to personality which take the interpersonal situation as a major focus of concern. Accordingly, most writers in the role-theoretical tradition (Cottrell, 1942; Gough, 1948; Mead, 1934; Sarbin, 1968) have given careful attention to this aspect of social functioning. Mead, for example, has argued that role-taking ability is the key variable in social and moral development; extending this line of reasoning he equates the "g" factor in intelligence with social sensitivity, the origins of which can be found in the central nervous system. In a similar vein, Cottrell and Dymond (1949) also maintained that empathy is the basic process in all social interaction.

Empathy, seen as an everyday manifestation of the disposition to adopt a broad moral perspective, to take "the moral point of view," also becomes important within the context of

research in moral development. By taking the moral point of view, a person is said to consider the consequences of his actions for the welfare of others (cf. Baier, 1958); in its social implications, this resembles Mead's notion of taking the role of the generalized other. A willingness or tendency to put oneself in another person's place and to modify one's behavior as a result, is clearly an important aspect of moral growth.

Because empathy plays a central role in many phases of social psychology, a valid and easily administered measure would be quite useful. However, most attempts to develop such a measure have not been particularly well received. Four examples might be mentioned.

Dymond (1948, 1949, 1950) developed a measure of empathic ability using rating forms. The results of Dymond's research have great intuitive appeal; unfortunately the method used to obtain empathy scores is cumbersome, and has been questioned on methodological grounds (Hastorf & Bender, 1952). The validity of the Empathy Test (Kerr, 1947) has been questioned by reviewers (cf. Hall, 1965; Thorndike, 1959). Taken at face value, the Test of Social Insight (Cassel, 1963) might seem promising. The consensus of reviewers (Black, 1965; Bordin, 1960) appears to be, however, that viewed more closely the test lacks validity and has little to do with an empathic or perceptive capacity for understanding others. More recently, new evidence has been presented for

¹ This paper is based on a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of California, Berkeley. The research was supported by Public Health Service Fellowship 1 F1 MH 30, 101-01A1 (MTLH), and was conducted at the Institute of Personality Assessment and Research, University of California, Berkeley. The author wishes to thank D. W. MacKinnon, Director of the Institute, for making its facilities available. Gratitude is also due K. H. Craik, H. G. Gough, and W. B. Hall for their generous advice and support.

² Requests for reprints should be sent to Robert Hogan, Department of Psychology, Johns Hopkins University, Baltimore, Maryland 21218.

the validity of the Chapin Social Insight Test (Chapin, 1942; Gough, 1965b). Although quite interesting and useful, the Chapin test seems more closely tied to the realm of social and interpersonal effectiveness than to the dimension of empathy per se, although overlap certainly exists.

In spite of the apparent difficulty involved in developing a valid and acceptable measure of empathy, the theoretical import of the concept requires that continuing efforts be made. The present report provides information concerning the origins and initial validation of another such attempt.

Is Empathy an Artificial Concept?

In view of the apparent gap between theory and measurement, the possibility exists that the term "empathy" is a creature of Academia. The practical appeal of the concept would be enhanced if intelligent laymen were found to use it in a meaningful fashion.

The consensus of dictionaries is that empathy means the intellectual or imaginative apprehension of another's condition or state of mind without actually experiencing that person's feelings. Although this definition implies a task, it says nothing about accuracy of performance. Empathy refers only to the act of constructing for oneself another person's mental state; the verisimilitude of the resulting construct is not a necessary part of the concept's meaning.

Fourteen nonpsychologists (school teachers, statistical clerks, etc.) were provided this definition of empathy. They were then asked to describe, using this definition, their conception of a highly empathic man with 50 items selected from the 100-item California Q sort (Block, 1961); these items were chosen for content directly relevant to empathy. The 14 individual Q-sort descriptions were intercorrelated, and the mean interjudge correlation was .53; the estimated reliability of the total composite from the Spearman-Brown correction was .94 (cf. Block, 1961, p. 37). This finding suggests that the group had a common behavioral referent for the concept of empathy and could use the term in a more or less consistent manner.

Similarly, to check agreement among psy-

chologists, nine graduate students in psychology were also asked to contribute a 50-item Q-sort description of an empathic man. Here the intercorrelations had a mean of .51; the estimated composite reliability was .90.

Finally, to determine the amount of agreement between laymen and psychologists, the two Q-sort composites were correlated, yielding a coefficient of .86; when corrected for attenuation in r_{11} and r_{22} , this becomes .93.

These findings suggest that people hold a common conception of the behavioral connotations of empathy. We should also inquire about the uniqueness of this conception, that is, to what degree will the composite Q sort for an "empathic man" correlate with a similarly derived description of "a good man," or "a mentally healthy man"?

Using the same 50 items, 10 undergraduate psychology majors were asked to contribute Q-sort descriptions of "a good man," where good was to be taken in a general moral or ethical sense. The 45 correlations between their descriptions had a mean of .62, the estimate reliability of the composite description of "a good man" was .94. The correlation between this composite and the psychology graduate students' "empathy" composite was .29, .32 when corrected for attenuation.

Next, 10 graduate students and faculty members in psychology at Johns Hopkins were asked to describe "the mentally healthy man," using the same Q sort as before. The average correlation of these descriptions was .65, and the reliability of the composite was estimated to be .95. The correlation between the composite description of the empathic and the mentally healthy man was .47, .51 corrected for attenuation. On the basis of role-theoretical considerations (i.e., an empathic disposition leads to effective social functioning which, in turn, facilitates mental health), we would expect this latter correlation to be moderately high, but certainly less than .9. The obtained correlation is considered both appropriate and conceptually meaningful. (Note: Each of the Q-sort composites discussed above was developed independently, thus there is no overlap in persons among the groups.)

Method

From these analyses, it can be inferred that empathy refers to a relatively discrete social phenomenon recognizable in the experience of laymen and psychologists alike. These findings serve as justification for the development of our empathy scale, and also provide a basis for its construction. The scale was assembled in the following manner. First, a criterion for assigning ratings of empathy was developed. Second, evidence was gathered for evaluating the ratings. Third, the ratings were used to define highly empathic and nonempathic groups. Fourth, an item analysis was conducted to select items for a scale which would discriminate between the nominated groups; the resulting scale might then serve as a measure of a general empathic disposition.

The empathy criterion was constructed as follows. Four faculty and research psychologists and three advanced graduate students in psychology at the University of California, Berkeley, were asked to describe their conceptions of a highly empathic man. They used the definition of empathy given above, and the full California *Q* sort (100 items). The seven *Q*-sort descriptions were intercorrelated, the coefficients ranged from .59 to .78, with a mean of .71; the estimated reliability of this composite was .94. This composite description of a highly empathic man served as our empathy criterion. To illustrate the criterion, the following are the five items selected as most characteristic of an empathic man (Block, 1961):

1. Is socially perceptive of a wide range of interpersonal cues.
2. Seems to be aware of the impression he makes on others.
3. Is skilled in social techniques of imaginative play, pretending, and humor.
4. Has insight into own motives and behavior.
5. Evaluates the motivation of others in interpreting situations.

The content of these items is clearly relevant to empathy: all reflect insight, perceptiveness, and social acuity. The five items chosen as least characteristic of a highly empathic man were:

1. Does not vary roles; relates to everyone in the same way.
2. Judges self and others in conventional terms like "popularity," "the correct thing to do," social pressures, etc.
3. Is uncomfortable with uncertainty and complexities.
4. Extrapunitive; tends to transfer or project blame.
5. Handles anxiety and conflicts by, in effect, refusing to recognize their presence; repressive or dissociative tendencies.

These items embody a complex dimension of conventionality, antiintreception, and reliance on the defensive techniques of projection and repression.

Since 1950 many different groups of people have been assessed at the University of California's Institute of Personality Assessment and Research (IPAR). Two groups of IPAR Ss were used in the present research. The first consisted of 100 military officers; the second sample contained 45 research scientists and 66 student engineers. Each of these Ss was studied by 8-10 skilled Os who recorded their impressions, formed during weekend living-in assessments, on the *Q* sort; composite *Q*-sort descriptions of these persons are on file at the Institute.

The composite *Q*-sort description of each individual in the two samples was correlated with the empathy criterion described above. The resulting correlation coefficient was considered the empathy rating for that person. The ratings ranged between $-.58$ and $+.68$ in the two samples.

Attempts to validate the criteria used in scale development lead logically to an infinite regress; thus, serious validation efforts are most appropriately directed at the measures themselves. Nonetheless, in view of the distinctly *a priori* nature of our empathy ratings, it seemed important to examine some of their behavioral and interpersonal implications.

The first indication of these implications comes from the following. Using the present sample of military officers, Gough (1955) developed a "Total Social Acuity Index" based on weighted scores derived from the Kerr Empathy Test, the Chapin Social Insight Test, and the Cline movie procedure (Cline, 1953). This index of social acuity correlated .26 ($p < .01$) with the military officers' empathy ratings.

Next, for the 45 research scientists and 40 of the student engineers, scores based on overall charades performance (Barron, 1954) correlated .61 and .58 with empathy ratings. These figures are to some degree inflated because many of the Os who contributed to the charades performance scores also provided *Q*-sort descriptions of the same Ss.

Finally, empathy ratings were correlated in each sample with the standard scales of the California Psychological Inventory (CPI; Gough, 1964), the MMPI (Hathaway & McKinley, 1943), and the Chapin Social Insight Test. Empathy ratings were only modestly related to these standard measures; the largest correlations were with the Capacity for Status (.24), Achievement via Independence (.22), Intellectual Efficiency (.22), Self-Acceptance (.21), and Achievement via Conformance (.20) scales of the CPI.³ In general, the correlations were positive for the CPI and the Chapin test which stress effective social functioning, and negative for the MMPI which has the opposite orientation. These correlations suggest that empathy ratings define a dimension which includes social competence, intellectual promise, and feelings of self-worth.

Our examination of empathy ratings permits two conclusions. First, use of the ratings as criterion mea-

³ In these samples a correlation of .20 is significant at the .05 level.

asures of empathy seems justified to some degree. Second, the ratings reflect an important aspect of social performance which is not measured well by a variety of widely used personality scales.

The empathy scale was developed by the standard technique of an item analysis of the responses of high-rated versus low-rated groups. From empathy ratings, Ss in the military officer sample, and in the sample of research scientists and student engineers, were placed into high (27%), middle (46%), and low (27%) subgroups. Studying each sample separately, the responses of the high and low subgroups were compared for 957 true-false items in the CPI, the MMPI, and an IPAR pool of items, using the chi-square or Fisher's exact statistic to evaluate differences. From these two analyses, 64 items (32 scored true, 32 false) were selected for the final scale.

Although the foundations of the empathy scale rest on a comparison of the responses of high- and low-rated groups, final selection of the items was not a "blind-empirical" procedure. Items for the final scale were retained on the basis of four considerations. The first was that differences in endorsement frequencies between the high- and low-rated groups be in the same direction in both samples (61 items). Second, McClelland (1951) performed an analysis of the MMPI item pool against a criterion of rated role-playing ability, and Gough (1955) item-analyzed the CPI-MMPI pool against his Social Acuity Index mentioned earlier. Fifteen items which appeared in these two earlier analyses also attained significance in the present research and were retained. Third, 17 of the items finally selected failed to attain statistical significance, but were retained on the basis of relevant content. Finally, items were chosen with an eye toward balancing the scale's true and false keying.

The following 64 items constitute the empathy scale. Thirty-one are from the CPI, 25 are from the MMPI, and the remaining eight items come from various experimental testing forms used in studies at IPAR in Berkeley:

A. *Scoring of items in the CPI:* 4(T), 8(T), 25(F), 52(T), 67(F), 79(F), 81(F), 84(T), 86(T), 97(T), 98(F), 100(T), 127(T), 186(F), 191(T), 194(F), 198(T), 239(T), 247(F), 255(F), 271(F), 275(T), 287(T), 359(T), 361(F), 363(F), 364(F), 403(T), 421(F), 442(F), 463(F).

B. *Scoring of items in the MMPI:* 15(F), 26(F), 57(T), 73(F), 78(T), 79(F), 100(T), 129(F), 170(F), 204(T), 231(T), 244(F), 248(T), 254(T), 327(F), 336(F), 355(T), 372(T), 399(T), 404(F), 407(T), 410(T), 417(F), 463(F), 478(F).

C. *Items from IPAR testing devices:*⁴

1. As a rule I have little difficulty in "putting myself into other people's shoes." (T)
2. I have seen some things so sad that I almost felt like crying. (T)
3. Disobedience to the government is never justified. (F)

4. It is the duty of a citizen to support his country, right or wrong. (F)
5. I am usually rather short-tempered with people who come around and bother me with foolish questions. (F)
6. I have a pretty clear idea of what I would try to impart to my students if I were a teacher. (T)
7. I enjoy the company of strong-willed people. (T)
8. I frequently undertake more than I can accomplish. (T)

Validity and Reliability of the Scale

We turn next to an examination of the validity and reliability of the empathy scale. Extensive analyses have been conducted with it, only a portion of which will be reported here. The first question to ask is, "How well does the scale meet its design requirements?" The empathy scale was built to predict Q-sort-derived empathy ratings. In the samples used in its development ($N = 211$), the average correlation between the scale and empathy ratings was .62; in an independent sample of medical school applicants ($N = 70$), this figure was .39.

A further check of the scale's validity is provided by the following: Five groups of Ss studied at IPAR were rated by the assessment staff for "social acuity," defined as:

The ability to respond intuitively and empathically to others and to group situations. A person high on this variable would be described as insightful, perceptive, and discerning, as having the knack of "sizing up" social situations and of making sound and dependable evaluations of people. He will be receptive to nuances and subtleties of behavior which might escape the notice of others, and will possess a flair for interpreting and integrating his observations.

The initial reliability of social acuity ratings for these samples ranged between .52 and .77, average .69; applying the Spearman-Brown formula to estimate the reliability of the composite ratings, the coefficients varied between .68 and .86, average .80. The mean correlation between empathy scale scores and rated social acuity in the samples used to develop the scale was .58. This figure is to some degree inflated because most of the Os who made social acuity ratings provided Q-sort descriptions of the same Ss. These Ss, in turn, were used to develop the scale. In an

⁴The author wishes to thank D. W. MacKinnon for permission to reproduce these items.

independent sample of 70 medical school applicants, the empathy scale correlated .42 with rated social acuity.

Some evidence also exists for the validity of the empathy scale when used with a younger population. The scale was administered to 121 junior high school students (51 boys and 70 girls) in the 13-15 year age range.⁵ Two teachers were given the definition of social acuity previously cited, then asked to rate the five most and five least socially acute boys in their respective classes. The procedure was repeated for the girls. Combining boys and girls separately, the empathy scale mean and standard deviation for the 10 "socially acute" boys defined in this manner were 33.0 and 4.1. Comparable values for the "nonsocially acute" boys were 27.2 and 4.3 ($t = 2.93$, $df = 18$, $p < .01$). For the 10 most and 10 least "socially acute" girls, the means and standard deviations were 36.2 and 5.3, 30.6 and 5.5, respectively ($t = 2.20$, $p = .05$).

With a sample of 50 college undergraduates, the reliability of the empathy scale estimated by a test-retest correlation after a 2-month interval, was .84. Applying the KR-21 formula to the scores of 100 military officers yields a coefficient of .71.

Psychological and Diagnostic Meaning

The empathy scale appears to have adequate reliability and concurrent validity to warrant further investigation of its psychological and diagnostic meaning.

If, as Mead (1934) suggested, empathic ability underlies social intelligence, then the empathy scale should relate to both measures of social effectiveness and functional intelligence. Table 1 provides evidence on this point. The empathy scale achieves its highest correlations with those scales from the CPI (1-5) specifically designed to measure social and interpersonal adequacy. The scale is also moderately related to Intellectual Efficiency, an index of the degree to which a person effectively mobilizes his intellectual resources.

The relationship between the empathy scale and intelligence scores is somewhat ambiguous.

⁵ Gratitude is expressed to Homer Jones, Principal of Vallejo Junior High School, for his cooperation in the testing of these students.

TABLE 1
CORRELATIONS BETWEEN EMPATHY SCALE
SCORES AND SELECTED VARIABLES

Tests and scales	Group	
	A	B
CPI		
1. Dominance	.48	.56
2. Capacity for Status	.62	.51
3. Sociability	.58	.56
4. Social Presence	.56	.60
5. Self-Acceptance	.34	.51
6. Well-Being	.19	.11
7. Responsibility	.11	.27
8. Socialization	.05	-.01
9. Self-Control	-.04	-.11
10. Tolerance	.42	.33
11. Good Impression	.34	.07
12. Communality	.00	.05
13. Achievement-Conformance	.30	.23
14. Achievement-Independence	.27	.17
15. Intellectual Efficiency	.52	.32
16. Psychological-Mindedness	.16	.46
17. Flexibility	.23	.29
18. Femininity	-.04	-.18
Scholastic Aptitude		
1. MCAT Verbal	-.16	
2. MCAT Quantitative	-.11	
3. SAT Verbal		.32
4. SAT Quantitative		.14
MMPI		
1. Lie Scale	.29	-.16
2. <i>F</i> (infrequent items)	-.15	-.25
3. <i>K</i> (ego functioning)	.35	.31
4. <i>Hs</i> (hypochondriasis)	.00	-.04
5. <i>D</i> (depression)	-.38	-.44
6. <i>Hy</i> (hysteria)	.17	.15
7. <i>Pd</i> (psychopathic deviate)	.11	.07
8. <i>Mf</i> (feminine interests)	-.01	.16
9. <i>Pa</i> (paranoia)	.15	-.01
10. <i>Pt</i> (psychasthenia)	-.21	-.05
11. <i>Sc</i> (schizophrenia)	-.11	-.07
12. <i>Ma</i> (mania)	.44	.14
13. <i>Si</i> (social introversion)	-.65	-.48
14. <i>A</i> (anxiety)	-.40	-.41
15. <i>R</i> (repression)	-.09	-.18
16. <i>Es</i> (ego strength)	.34	.38
17. <i>Sd</i> (social desirability)	.50	.37
18. <i>TMA</i> (manifest anxiety)	-.49	-.20

Note.—Group A: 70 medical school applicants $r = .23$, $p < .05$; $r = .30$, $p < .01$. Group B: 51 female college seniors $r = .27$, $p < .05$; $r = .35$, $p < .01$. MCAT: Medical College Aptitude Test.

ous. The correlations presented between the scale, the MCAT, and the SAT in Table 1 show this ambiguity—coefficients ranging from low negative to moderately positive

depending on the population and the measure of intelligence used.

The empathy scale has been compared with a number of other personality assessment devices to gain broader understanding of its predictive implications. Correlations with MMPI scales are also given in Table 1. From these it can be inferred that high scorers on empathy will be sociable, optimistic, and free from unnecessary doubts and worries.

Table 2 offers comparisons between the empathy scale and indexes from the Myers-Briggs Type Indicator (Myers, 1962) and the Study of Values (Allport, Vernon, & Lindzey, 1960).

Certain single personality scales have been so widely used in research that any report of a new measure must indicate its relationship to them. Table 1 provides correlations between the empathy scale, the Ego Strength scale (Barron, 1955), the Manifest Anxiety scale (Taylor, 1953), and Edwards' scale for Social Desirability (Edwards, 1957). Table 2 gives evidence concerning the relationship between the empathy scale and Eysenck's

(1959) factor-pure measures of neuroticism and extraversion.

Next, we must consider the empathy scale's relation to both authoritarian tendencies and acquiescence, for which the California F scale (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950; Bass, 1955) is an accepted index. In a sample of 48 male undergraduates the correlation between the Empathy scale and the F scale was $-.52$; the same index in a sample of 29 female undergraduates was $-.30$. Finally, in the same sample of undergraduate men, the empathy scale correlated $-.31$ with the Dogmatism scale (Rokeach, 1960).

An examination of the correlations between the empathy scale and existing personality measures shows several moderately large coefficients. Given the size of certain of these correlations, it might be argued that the empathy scale is only a redundant measure of, for example, extraversion. Such a proposal would overlook two important points. First, the empathy scale predicts such diverse phenomena as low anxiety, self-acceptance, and lack of authoritarian tendencies as well as it does extraversion. Thus one may also argue that the empathy scale reflects attributes common to a wide range of interpersonal behavior rather than redundantly measuring one specific aspect of social functioning. A more fundamental indication of the scale's utility, however, is that it predicts rated empathy better than all existing measures with which comparisons have been made.

The empathy scale has been scored for 11 groups of men and three groups of women (total $N = 1,086$). Means and standard deviations for these groups are presented in Table 3. The highest mean score was obtained by a small group of upper-division men enrolled in a personality assessment course at the University of California. The class contained for the most part students hoping to do graduate work in clinical psychology. The "Education abroad students" were highly selected students chosen to participate in the University's overseas program. Means for the two delinquent groups have been estimated by scoring the 39 empathy scale items found in the CPI. In two comparisons ($N = 170$), this CPI empathy subscale cor-

TABLE 2

CORRELATIONS BETWEEN EMPATHY SCALE SCORES
AND THE VARIABLES INDICATED

Tests and scales	r
Study of Values ^a	
1. Theoretical	$-.15$
2. Economic	$.17$
3. Aesthetic	$.08$
4. Social	$.19$
5. Political	$.00$
6. Religious	$-.14$
Maudsley Personality Inventory ^b	
1. Neuroticism	$-.28$
2. Extraversion	$.51$
Myers-Briggs Type Indicator ^c	
1. Extraversion	$.63$
2. Introversion	$-.61$
3. Sensation	$-.16$
4. Intuition	$.16$
5. Thinking	$-.18$
6. Feeling	$.10$
7. Judging Function	$-.17$
8. Perceptive Function	$.10$

^a $N = 43$ males, $r = .30$, $p < .05$.

^b $N = 60$ males, $r = .25$, $p < .05$; $r = .33$, $p < .01$.

^c $N = 42$ males, $r = .30$, $p < .05$; $r = .39$, $p < .01$.

TABLE 3
NORMATIVE DATA FOR THE EMPATHY SCALE
ON THE SAMPLES INDICATED

Group	N	M	SD
Males			
1. Psychology majors	24	44.7	5.2
2. Education abroad students	18	43.2	4.8
3. Medical students	70	42.4	5.3
4. Research scientists	45	40.3	5.4
5. College students (all levels)	90	39.1	4.7
6. Honor students in engineering	66	38.0	6.4
7. Military officers	100	37.7	7.1
8. Architects	124	36.8	5.3
9. Junior high school students (Grades 7 and 8)	51	31.0	5.0
10. Prison inmates	92	30.4	—
11. Young delinquents	100	29.1	—
Females			
1. College seniors	143	41.5	5.1
2. College students (all levels)	93	40.7	5.9
3. Junior high school students (Grades 7 and 8)	70	33.7	5.2

related .92 with the full-scale score. Additionally, in these groups the mean of the item p values for the 39 items is .44, the mean p value for the remaining 25 items is .50. Thus the prorated means underestimate, but only very slightly, the appropriate full-scale values. It should also be noted that females score one to two points higher than males on the average; that is, women seem to be slightly more empathic than men—a point which accords well with conventional wisdom on this topic.

Interpersonal Implications and Psychological Basis of Measurement

Gough (1965a) has presented a strategy for discovering and analyzing the psychological meaning of test scores. The method proceeds through three stages. In the first we ask how well a scale relates to its appropriate criteria. The second stage rests on analyses intended to reveal the scale's underlying basis of measurement. The third requires a search for unusual or unexpected relationships. From Gough's scheme, data presented thus far are necessary for the primary evaluation of a scale. The evidence indicates that the empathy scale predicts to some degree what it seeks

to predict. It is appropriate therefore to move to the second level of evaluation, that is, exploration of the scale's psychological basis of measurement and personological meaning (insufficient evidence is available to permit discoveries at the third level).

A fundamental question at the second level of analysis is, "Does the scale meet the requirements set by its conceptual base?" The empathy scale was developed within the framework of a multidimensional theory of moral development (Hogan, 1967). In contrast with psychoanalysis, for example, the theory predicts that a person who fails to internalize parental prohibitions and the rules of society will not necessarily become delinquent. Rather, if he learns to recognize the existence of the rights of others, if he can adopt the moral point of view, then he may behave appropriately despite experiencing society's rules as alien and external. Internalizing social prohibitions and learning to take the moral point of view are seen as two independent stages in moral development.

From this discussion we predict that there will be differences in the behavior of poorly socialized groups differing in the degree to which they are disposed to take the moral viewpoint. To test this expectation, the Socialization(So) scale of the CPI was selected as an appropriate index of strength of internalization of social rules—evidence supporting this choice is substantial (cf. Gough, 1965a). The empathy scale was assumed to reflect the ability to adopt the moral point of view. In a sample of 429 nondelinquent Ss, 124 (40.3% of the sample) had low So scores (raw score < 33). From this latter group, 50 Ss were chosen, of whom 25 had low scores on the empathy scale (raw score < 30), and 25 had high scores (raw score > 44). Persons with low scores for both socialization and empathy were placed in one group, low-socialization-high-empathy Ss were placed in a second group. Using the CPI, a scale by scale comparison of the two groups was conducted, the results of which appear in Table 4. The two groups differed significantly on 12 of the 17 scales which could be meaningfully compared. When the mean scores for the two groups are plotted on a CPI profile sheet, the resulting profiles are quite distinct-

TABLE 4

COMPARISON OF HIGH-EMPATHY-LOW-SOCIALIZATION AND LOW-EMPATHY-LOW-SOCIALIZATION GROUPS ON CPI SCALES

Scale	Group A		Group B		<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
1. Dominance	32.2	4.5	26.9	6.1	3.5	.01
2. Capacity for Status	23.6	2.9	19.2	4.0	4.5	.01
3. Sociability	26.6	4.1	24.0	6.2	1.7	.10
4. Social Presence	42.6	4.8	35.4	7.2	4.7	.01
5. Self-Acceptance	25.0	2.6	22.4	6.6	1.8	.10
6. Well-Being	37.4	3.8	34.0	4.6	2.9	.01
7. Responsibility	29.4	4.5	27.3	4.6	1.6	<i>ns</i>
8. Socialization ^a	28.8	2.4	29.4	4.1	-0.7	<i>ns</i>
9. Self-Control	25.4	6.0	23.1	7.2	1.2	<i>ns</i>
10. Tolerance	25.0	4.8	20.6	4.8	3.2	.01
11. Good Impression	16.2	3.8	13.4	6.2	1.9	.10
12. Communality	25.0	2.0	25.5	3.2	-0.7	<i>ns</i>
13. Achievement-Conformance	28.2	4.0	23.6	3.9	4.1	.01
14. Achievement-Independence	24.1	3.2	21.8	6.3	1.6	<i>ns</i>
15. Intellectual Efficiency	43.0	4.1	35.5	7.9	4.3	.01
16. Psychological-Mindedness	15.2	3.6	12.1	3.5	3.1	.01
17. Flexibility	14.6	3.7	12.4	4.8	1.8	.10
18. Femininity	16.5	3.0	14.8	5.1	1.5	<i>ns</i>
Empathy Scale	46.4	2.7	29.4	3.8	18.1	.01

Note.—Group A: high empathy, low So; *N* = 25. Group B: low empathy, low So; *N* = 25.
^a Scale on which groups were selected for low scores.

tive. The low-So-high-empathy profile is favorable. The group characterized by low scores on both So and empathy appears marginally adjusted and prone to antisocial behavior. These observations, based on inferences from the profiles, give tentative support to the conceptual meaning of the empathy scale.

Extending the implications of the analysis given above, if the empathy scale reflects the disposition to adopt the moral point of view, then legal offenders should receive lower scores than nondelinquents. To test this prediction, the empathy scale was applied to a group of 92 inmates at a state prison. In this case only the 39 empathy scale items appearing on the CPI were scored. The prisoners' scores on the Army General Classification Test ranged between the 15th and the 99th percentile. Their mean was at the 52nd percentile, indicating that the group was of approximately average intelligence. For the prisoners, the mean and standard deviation on the 39 item empathy subscale were 18.5 and 5.1. These values can be compared with

those from a group of 100 military officers, which were 22.7 and 4.8 on the same scale. The difference in the means of these groups is highly significant ($p < .001$). Using Tilton's (1937) method as an estimate, there is 39.5% overlap in the scores of the two groups. Although the officers are undoubtedly more intelligent than the prisoners, it is doubtful that the obtained differences in empathy scores can be explained entirely in terms of differences in intellectual endowment.

A sample of 103 fraternity men was chosen to perform a final empirical check on the interpersonal meaning of the scale. The empathy scale mean and standard deviation for this group was 38.6 and 6.2. Each member had also been described by five of his peers on the 300-word Gough Adjective Check List (Gough, 1960; Gough & Heilbrun, 1965). A composite description of each person can be obtained from these five individual descriptions by counting the number of times each word is checked, and using the totals as scores. Thus, if Jones is checked by all five of his peers as "friendly," his score

on this adjective will be 5. Jones will have then 300 scores, one for each adjective.

From 33 adjectives with correlations significant at or above the .05 level, those 10 with the most positive, and the 10 adjectives with the most negative correlations are listed below.

The 10 words most descriptive of high scorers and the correlations are:

pleasant	(.32)	sociable	(.26)
charming	(.28)	sentimental	(.26)
friendly	(.27)	imaginative	(.26)
dreamy	(.27)	discreet	(.25)
cheerful	(.26)	tactful	(.24)

The high scorer on the empathy scale seems likable and friendly, possessing considerable charm, poise, and tact. He is outgoing, warm, and very much at ease in the interpersonal situation—the term urbane might even be appropriate.

The 10 words most descriptive of low scorers are:

cruel	(−.37)	unemotional	(−.23)
cold	(−.29)	unkind	(−.23)
quarrelsome	(−.27)	hard-hearted	(−.23)
hostile	(−.24)	argumentative	(−.21)
bitter	(−.24)	opinionated	(−.21)

Thus the low scorer on the empathy scale appears somewhat aloof, disaffected, and disposed to alienate those around him.

Discussion

A proper evaluation of the empathy scale must answer two questions. First, "To what degree do scale scores reflect an empathic disposition?" This question is obviously directed toward the primary validity of the scale itself. The second question, "Do highly empathic persons, as identified by the scale, adopt the moral point of view more often than non-empathic persons?" has implications for both the scale and the model of moral development from which it was derived.

In reply to the first question, the empathy scale appears to work well. From the peer ratings, test correlates, and other validation evidence presented above, high scorers seem to be socially acute and sensitive to nuances in interpersonal behavior. Low scor-

ers, on the other hand, seem hostile, cold, and insensitive to the feelings of others.

The second question, "Do highly empathic persons adopt the moral point of view more often than nonempathic persons?" is harder to answer. This is due in part to the highly inferential nature of the judgment involved; that is, behavior which entails adopting the moral point of view is often hard to recognize. Additionally, it is always difficult to answer adequately questions about the construct validity of a personality scale. It would appear that delinquents score lower on the scale than nondelinquents. Furthermore, low-empathy scores seem to have negative implications for the social conduct of persons in an otherwise normal population. We conclude that empathic ability, as measured by the empathy scale, seems to be at least one requirement for taking the moral point of view. Obviously more research is required for a complete answer to this question.

It seems fair to conclude at this point that both the practical and theoretical value of the empathy scale have received a measure of support. However, the evidence presented here represents only a first step toward a truly definitive evaluation. It is hoped that the potential merits of the scale will attract the attention of psychological researchers so that its utility can be properly assessed.

REFERENCES

- ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality*. New York: Harper, 1950.
- ALLPORT, G. W., VERNON, P. E., & LINDZEY, G. *Manual: Study of values*. (3rd ed.) Boston: Houghton Mifflin, 1951.
- BAIER, K. *The moral point of view*. New York: Random House, 1958.
- BARRON, F. *Charades as an assessment device*. Technical Memorandum, November, 1954, Project No. 7730, Contract No. AF 18(600)-8, Officer Education Research Laboratory, Air Force Personnel and Training Research Center, Maxwell Air Force Base, Alabama.
- BARRON, F. An ego-strength scale which predicts response to psychotherapy. *Journal of Consulting Psychology*, 1955, 17, 327-333.
- BASS, B. M. Authoritarianism or acquiescence? *Journal of Abnormal and Social Psychology*, 1955, 51, 616-623.
- BLACK, J. D. Test of social insight. Reviewed in O. K. Buros, Jr. (Ed.), *Sixth mental measurements*

- yearbook. Highland Park, N. J.: Gryphon Press, 1965.
- BLOCK, J. *The Q-sort method in personality assessment and psychiatric research*. Springfield, Ill.: Charles C Thomas, 1961.
- BORDIN, E. S. Test of social insight. Reviewed in *Journal of Consulting Psychology*, 1960, 24, 100.
- CASSELL, R. N. *Test of social insight*. (Rev. ed.) New Rochelle, N. Y.: Bruce, 1963.
- CHAPIN, F. S. Preliminary standardization of a social insight scale. *American Sociological Review*, 1942, 7, 214-225.
- CLINE, V. B. The assessment of good and poor judges of personality using a stress interview and sound film technique. Unpublished doctoral dissertation, University of California, Berkeley, 1953.
- COTTRELL, L. S. The analysis of situational fields in social psychology. *American Sociological Review*, 1942, 7, 370-382.
- COTTRELL, L. S., & DYMOND, R. A. The empathic responses: A neglected field for research. *Psychiatry*, 1949, 12, 355-359.
- DYMOND, R. A. A preliminary investigation of the relation of insight and empathy. *Journal of Consulting Psychology*, 1948, 12, 228-233.
- DYMOND, R. A. A scale for the measurement of empathic ability. *Journal of Consulting Psychology*, 1949, 13, 127-133.
- DYMOND, R. A. Personality and empathy. *Journal of Consulting Psychology*, 1950, 14, 343-350.
- EDWARDS, A. L. *The social desirability variable in personality assessment and research*. New York: Dryden, 1957.
- EYSENCK, H. J. *The Maudsley Personality Inventory*. London: University of London Press, 1959.
- GOUGH, H. G. A sociological theory of psychopathy. *American Journal of Sociology*, 1948, 53, 359-366.
- GOUGH, H. G. *The assessment of social acuity*. Research Report, March, 1955, Project No. 7730, Contract No. AF 18(600)-8, Officer Education Research Laboratory, Air Force Personnel and Training Research Center, Maxwell Air Force Base, Alabama.
- GOUGH, H. G. The adjective check list as a personality assessment research technique. *Psychological Reports*, 1960, 6, 107-122.
- GOUGH, H. G. *Manual for the California Psychological Inventory*. (Rev. ed.) Palo Alto, Calif.: Consulting Psychologists Press, 1964.
- GOUGH, H. G. Conceptual analysis of psychological test scores and other diagnostic variables. *Journal of Abnormal Psychology*, 1965, 70, 294-302. (a)
- GOUGH, H. G. A validation study of the Chapin social insight test. *Psychological Reports*, 1965, 17, 355-368. (b)
- GOUGH, H. G., & HEILBRUN, A. B., JR. *Manual for the adjective check list*. Palo Alto, Calif.: Consulting Psychologists Press, 1965.
- HALL, W. B. The empathy test. Reviewed in O. K. Buros, Jr. (Ed.), *Sixth mental measurements yearbook*. Highland Park, N. J.: Gryphon Press, 1965.
- HASTORF, A. H., & BENDER, I. E. A caution respecting the measurement of empathic ability. *Journal of Abnormal and Social Psychology*, 1952, 47, 574-576.
- HATHAWAY, S. R., & MCKINLEY, J. C. *Manual for the Minnesota Multiphasic Personality Inventory*. New York: Psychological Corporation, 1943.
- HOGAN, R. T. Moral development: An assessment approach. Unpublished doctoral dissertation, University of California, Berkeley, 1967.
- KERR, W. A. *The empathy test*. Chicago: Psychometric Associates, 1947.
- MCCLELLAND, W. A. A preliminary test of role-playing ability. *Journal of Consulting Psychology*, 1951, 15, 102-108.
- MEAD, G. H. *Mind, self, and society*. Chicago: University of Chicago Press, 1934.
- MYERS, I. B. *Manual for the Myers-Briggs Type Indicator*. Princeton: Educational Testing Service, 1962.
- ROKEACH, M. *The open and closed mind*. New York: Basic Books, 1960.
- SARBEN, T. R., & ALLEN, V. L. Role theory. In G. Lindzey & E. Aronson, *Handbook of social psychology*. Vol. 1. (2nd ed.) Reading, Mass.: Addison-Wesley, 1968.
- TAYLOR, J. A. A personality scale of manifest anxiety. *Journal of Abnormal and Social Psychology*, 1953, 48, 285-290.
- THORNDIKE, R. L. The empathy test. Reviewed in O. K. Buros, Jr. (Ed.), *Fifth mental measurements yearbook*. Highland Park, N. J.: Gryphon Press, 1959.
- TILTON, J. W. The measurement of overlapping. *Journal of Educational Psychology*, 1937, 28, 656-662.

(Received June 14, 1968)

PERSONALITY CHARACTERISTICS OF THERAPISTS: DESCRIPTION OF RELEVANT VARIABLES AND EXAMINATION OF CONSCIOUS PREFERENCES¹

ARON Z. SPILKEN,² MARTIN A. JACOBS, JAMES J. MULLER
AND JANE KNITZER

Boston University School of Medicine

Recent studies of the significance of the psychotherapist's personality are summarized and discussed with the aim of developing measures for clinical research. Ten variables are described, comprising a relatively comprehensive and discrete list of potential importance. They include: objectivity, dependability, sincerity, sureness, directiveness, empathy (both affective and cognitive components, i.e., responsiveness and understanding), respect, interest, and warmth. The measurement of conscious attitudes toward these variables in 54 psychiatrists, 77 psychologists, and 78 social workers is described. The method uses a questionnaire requiring forced choices between paired adjectives and provides rankings among the variables. The results indicate both similarities and differences among therapist groups in relative preferences for these characteristics.

The process of psychotherapy, as revealed in the actual interaction between therapist and patient, reflects certain background characteristics of the therapist: his formal professional training, his years of experience, and his personal style. Increasing attention is currently being paid to the role of the last factor in determining whether or not treatment will be successful. In the past, most studies in this area have focused on a limited view of the therapist's personality, or else have described such characteristics in global terms, thereby failing to identify the various discrete elements inherent in the general categories.

This article is a report of an attempt to conceptualize these specific elements so that a full range of personal characteristics of the therapist relevant to the treatment process may be identified. To illustrate this approach,

some findings will be presented which examine the self-reported relative preferences assigned to these characteristics by different professional groups with varying levels of therapeutic experience.

Prominent among previous studies in this area has been the work of Whitehorn and Betz (1960), who used response patterns to the Strong Vocational Interest Blank as a means of differentiating successful from unsuccessful therapists of schizophrenic patients. Therapists designated Type A showed a resemblance in their responding to such vocational groups as lawyers and CPAs with their interest in verbal matters. Type B therapists bore a similarity to printers and math-science teachers with their interest in mechanical matters. These differences have been assumed to reflect variations in degree of understanding, permissiveness with patients, and ability to establish a trusting relationship.

Strupp, Wallach, and Wogan (1964), in a retrospective study of the psychotherapeutic process as experienced by the patient, observed that patients who reported having the therapist's respect and interest tended most often both to judge themselves and to be judged by their therapists as having been in successful treatment. Also contributing to a better outcome was the patient's experience of the therapist as warm, natural, unstudied,

¹ The results of this paper were discussed in part at the meeting of the American Psychological Association, San Francisco, September 1968. The authors are grateful for the cooperation of the individuals in the various professional groups who returned the completed forms, for the assistance of Chester C. Bennett, L. D. Carter, and James Mann in contacting the student (trainee) groups, and for the statistical and clerical help of Sylvia Fleisch, Phyllis Garbose, Carol Lindop, and Patricia Batson.

² Requests for reprints should be sent to Aron Z. Spilken, Boston University School of Medicine, Division of Psychiatry, 80 East Concord St., Boston, Massachusetts 02118.

and attentive, and not as cold, distant, formal, or extremely passive. Patients who were uncertain about how the therapist felt toward them tended to have a less successful treatment experience.

Evaluating therapy behavior by training orientation, Sundland and Barker (1962) surveyed 139 psychotherapists. They found that a general factor (analytic-experiential) significantly distinguished among therapists identifying themselves with different schools (Freudian, Sullivanian, and Rogerian). The 13 scales used to tap this dimension included Frequency of Therapist's Activity, Type of Therapist's Activity, Emotional Tenor of the Relationship, Structure of the Relationship, and Therapist's Security.

McNair and Lorr (1964) analyzed the self-reports of 265 professional workers in a Veterans Administration setting. Psychiatrists, psychologists, and social workers indicated their preferred treatment techniques, along three dimensions derived from Sundland and Barker's (1962) work, which were felt by the authors to be critical in describing the therapeutic process. The three factors were analytic, impersonal, and directive. Analytic referred to the use of traditional psychoanalytic techniques, such as analysis of resistance and transference, and exploration of early childhood experiences. Impersonal referred to the therapist's preference for remaining reserved and uninvolved, as opposed to personalizing the relationship. Directive referred to a preference for planning a therapeutic strategy and guiding the therapeutic interaction in contrast to allowing the patient more freedom in shaping the course of treatment. They found that preferences for different technique patterns tended to distinguish among the professions, but were not a function of experience.

In contrast to the above studies based on patients' and therapists' reports, Truax and Carkhuff (1965) examined the therapeutic interaction directly through the analysis of tape recordings, focusing on therapist variables of self-congruence or genuineness, accurate empathy, and unconditional positive regard. Increases in the level of therapist empathy and positive regard within a single session were found to be related to depth of the patient's intrapersonal exploration. It was

further demonstrated (Truax, Wargo, Frank, Imber, Battle, Hoehn-Saric, Nash, & Stone, 1966) that therapists providing high levels of these conditions showed greater patient improvement than therapists providing lower levels.

These previous studies employed quite different methods and focused on limited groups of variables which overlapped conceptually. The various factors isolated in therapists' reports, patients' judgments, and direct observations should now be related in a single, comprehensive schema, describing the range of therapists' behavior. Further, the discrete elements inherent in global concepts should be identified in order to assess both their individual effects and their impact in various combinations.

Consistent with this aim, 10 discrete variables, derived both from prior studies and from clinical judgments were identified as relevant to the therapeutic process. Each variable was conceptualized as representing a continuum with the midpoint reflecting a moderate amount of a therapist characteristic and with high and low deviations at the extremes. The 10 variables are objectivity, dependability, sincerity, sureness, directiveness, responsiveness, understanding, respect, interest, and warmth.

1. *Objectivity.* This factor is defined as a moderate degree of involvement consonant with balanced perspective and freedom from personal bias. Very low objectivity reflects overidentification with the patient, extreme closeness, and loss of perspective. Very high objectivity reflects aloofness and detachment. McNair and Lorr (1964) dealt with similar endpoints in their interpersonal dimension, without providing for the midpoint. Similarly, Strupp et al. (1964) dealt with distance and formality as distinct qualities, rather than as points on a continuum.

2. *Dependability.* It may be expected that the degree to which the therapist is consistent and reliable largely determines the establishment of a trusting therapeutic relationship. At the midpoint of this scale the therapist indicates he is responsible and can be counted on to do what he says. The high end of the scale describes conscientiousness to the point of self-sacrifice and martyrdom on the part

of the therapist. At the low end, the therapist is irresponsible and unreliable.

3. *Sincerity*. Sincerity is derived from Truax's concept of genuineness, and Rogers' concept of therapist self-congruence. This refers to the presentation of a "real" person in the therapeutic interaction, and the degree of congruence between what the therapist is feeling at the moment and what he says. Whitehorn and Betz (1960) found that physicians effective with schizophrenic patients expressed personal attitudes more freely than less effective therapists. Truax rates this quality from low to high, making the assumption that the higher score is preferable. On our scale, we have allowed for the possibility of what might be considered by some as an inappropriately high degree of genuineness.

This factor in its moderate form is defined as truthfulness consistent with protection of patient's feelings. The endpoints of the scale reflect extreme openness, personal confession, and self-revelation at the high end, and superficiality, deception, or therapist role-playing at the low end.

4. *Sureness*. It may be expected that when a therapist is unsure, anxious, or lacking in confidence, therapeutic progress is disturbed. Knupfer, Jackson, and Krieger (1959) found an association between therapists' confidence in themselves and their level of competence. Bandura (1956) also sees freedom from anxiety as basic to competence and feels that the therapist's ability to respond nonanxiously is one of the crucial factors in his ability to relieve patient's anxieties. Excessive sureness, at the high end of this scale, is exemplified by the therapist manifesting grandiosity and appearing pompous and dogmatic. At the low end, the therapist seems anxious, inadequate, helpless, and indecisive. Moderation reflects the therapist appearing to believe in and know what he is doing; he appears at ease, comfortable, and confident.

5. *Directiveness*. Moderation at the midpoint, reflects the therapist who appropriately provides structure, guidance, and limit setting. At the high end of the scale the therapist is conceptualized as authoritarian, controlling, and manipulative; the patient is told what to do and what not to do both in and out of treatment. At the low end of the scale the

therapist adopts a passive, permissive stance, allowing the patient under all circumstances to proceed at his own rate and behave as he pleases. In the McNair and Lorr (1964) study, the authors found that directiveness differentiated among the therapeutic professions.

Empathy is the variable perhaps most universally accepted as vital to successful psychotherapy. It has been typically treated as a global quality describing a meaningful flow of understanding and communication between therapist and patient. Truax (1963), for instance, following the theoretical orientation of Rogers, designed a compound Accurate Empathy Scale to measure both the therapist's sensitivity to the patient's current feelings and his ability to communicate his understanding of these feelings in a way acceptable to the patient. Fox and Golden (1964) in summarizing the analytic literature point out that empathy has been conceptualized as both an emotional and intellectual process. To deal more precisely with empathy and to make it applicable to a wider range of theoretical approaches, it seemed desirable to separate this concept into its components of responsiveness and understanding, and to enlarge the concept of understanding to include the therapist's conceptualization of the role of feelings and defenses in the patient.

6. *Empathy (Responsiveness)*. At the high end of this scale the therapist is seen as sympathetic; he feels for and with the patient and communicates this to him. At the low end, the therapist may have intellectual knowledge, but no real emotional appreciation of the patient's feelings. At the midpoint the therapist has an emotional awareness of how the patient feels and communicates his alliance with the patient, but is not totally identified with him.

7. *Empathy (Understanding)*. At the high end of this scale, the therapist imposes his understanding before the patient is ready to accept it; that is, his interpretation is too early or too deep. At the low end, the therapist is insensitive to patient cues and lacks understanding, or is only partially aware of what the patient is trying to say. At the midpoint, the therapist comprehends implicit or latent meanings, and is able to place them

in context. His communications are both accurate and appropriately timed. This factor seems most related to the concept of therapeutic skill (Bergin, 1966).

8. *Respect*. At the midpoint of a discrete respect continuum, the therapist shows appreciation and acceptance of the patient's individual worth and uniqueness. At the high end, the therapist evidences uncritical or naïve acceptance of the patient's reports, without adequately assessing the facts. At the low end, the therapist does not appear to value the patient as an individual and may be judgmental, sarcastic, or insulting. Strupp et al. (1964) found that if the patient perceived the therapist as having respect for him, he was likely to report positive outcome. Others, such as Rogers (1962), seem to have incorporated respect into more generalized concepts. Unconditional positive regard, for example, seems to include warmth along with a respecting acceptance of the patient.

9. *Interest*. At the high end of this dimension, the therapist appears to be prying, overly curious, and intrusive. At the low end he seems inattentive and disinterested in the patient. At the midpoint he is attentive and alert to the patient's needs and communications.

10. *Warmth*. At the midpoint of this continuum, the therapist feels comfortable with and likes the patient, is reassuring and encouraging. At the high end, the therapist is infantilizing, fosters dependency which might be difficult to terminate, or promotes intimacy that is uncomfortable or embarrassing to the patient. At the low end, the therapist expresses no liking or affection for the patient and seems cold and unresponsive.

Thus, the authors' conception of therapists' personal and stylistic characteristics which are expected to influence the course of treatment includes these 10 variables. As a first step in experimentally examining these dimensions, it was decided to determine their relative value in the therapeutic community, comparing professed attitudes among therapists of different schools of training (psychiatry, psychology, and social work) and among therapists of different levels of experience. These professed values can then serve as a basis for comparisons in future studies where these character-

istics are observed in actual therapeutic behavior. The general assumptions were that these qualities are differentially valued by members of the various professions, and that experienced therapists feel differently about them than do the trainees.

METHOD

A scale was designed to determine the relative value which a therapist professed to have for each characteristic. Because self-ratings were to be obtained, only the midpoints of each characteristic could be expressed since many of the endpoints are clearly undesirable and would not be likely to be chosen as preferred traits. A list of synonyms representing each variable was prepared. Examples of these are: 1. objective, impartial; 2. dependable, reliable; 3. genuine, sincere; 4. confident, secure; 5. directing, guiding; 6. and 7. empathic, understanding; 8. accepting, respecting; 9. interested, attentive; 10. warm, encouraging. It will be noted that for purposes of this technique, responsiveness and understanding were treated as one variable (Concepts 6 and 7). Thus nine concepts were evaluated.

An adjective reflecting each of the nine variables was paired with adjectives of every other factor, resulting in 72 paired items. Each therapist was asked to select one term from each pair representing "qualities which are important in a good therapist," as of more value to him than the other. It was indicated, in the instructions, that this was difficult since all items were selected for their basic value in therapy. To accomplish this, however, the rater was asked to make a forced choice in each pair on the basis of his first impressions. The choice was structured as a "relative preference" rather than an absolute one. To assess the consistency of the ratings the therapist was asked to complete a parallel form approximately 24 hours later, without referring back to the first. The parallel form was essentially the same with different combinations of 72 paired synonyms. In order to obtain as truthful a reaction as possible, the raters were asked that the forms be completed anonymously and that they be returned to us in stamped, addressed envelopes which were coded to identify subsamples.

The sample tested consisted of experienced (clinic or hospital staff) and inexperienced (graduate students and residents) members of the therapeutic professions: psychiatry, clinical psychology, and social work. The vast majority of these were trained in the psychoanalytic orientation. Because it was possible that the values revealed might merely reflect professional status or education in general, as a control for this, dental and law students (i.e., individuals with postgraduate education) were also sampled.

Scoring of the Psychotherapist's Preference Scale (PPS) was accomplished by summing the number

of adjectives underlined for each category. The sum of choices on both Forms A and B was used to represent the best estimate of a rater's preferences. The scores for each of the nine concepts were then compared statistically across groups, both by profession and by experience, in order to determine whether the different samples of therapists attached greater or lesser value to the various characteristics. As a follow-up to the two-way analysis of variance, Duncan's multiple-range test (see Edwards, 1966) was used to determine precisely which mean scores were significantly different among the six therapist groups. In order to determine whether all professionals value these concepts equally, the 209 psychotherapists were compared directly with the 66 others in the sample by means of a one-way F test. Finally, in order to obtain an understanding of the relative importance attached to each concept by each class of raters, the preferences of each group were rank ordered, and multiple-range tests performed.

RESULTS

Table 1 indicates the number of completed forms returned by each group and the percentage this represented of total forms distributed to members of that group. Two hundred and nine psychotherapists were sampled, with an average rate of return of 71%. This suggests that the sample is representative of the population of psychotherapists. The rate of return for the 66 law and dental students was much lower (42%) indicating less concern with the study and less representativeness.

Reliability figures also appear in Table 1 and indicate that for psychotherapists in general consistent responses were given to both forms. Responses to the two forms were reliably in agreement for 86% of these Ss. The average rho was .78 for this subsample. The dentists and lawyers showed less consistency from Form A to B. Only 55% of this group was in reliable agreement. The average rho was .61 for this subsample. These differences between level of agreement are statistically significant ($\chi^2 = 27.83$, $p < .001$).

In order to determine whether all groups with postgraduate education equally value certain characteristics, psychotherapists and nonpsychotherapy professionals were compared on each individual concept. For all but one of the concepts there are highly significant differences between the scores of the two groups. Psychotherapists valued empathy, re-

TABLE 1
INFORMATION CONCERNING NUMBER AND RATE OF
RETURN AND RELIABILITY OF
COMPLETED FORMS

Group	Number return- ing com- pleted forms	Percent- age of total sampled	Percent- age evi- dencing signifi- cant agree- ment ^a	Average degree of cor- relation (rho)
Social workers	20	95	80	.79
Social work stu- dents	58	67	88	.75
Clinical psycho- logists	50	65	86	.79
Psychology trainees	27	79	96	.84
Psychiatrists	26	70	85	.77
Psychiatric residents	28	74	79	.76
Dental students	25	53	64	.62
Law students	41	37	49	.60
All therapists	209	71	86	.78
Dental and law students	66	42	55	.61

^a Percentage of Ss in a group achieving rho of at least .60, which is significant at the .05 level of confidence for nine paired concepts.

spect, and interest more highly, and objectivity, dependability, sincerity, sureness, and directiveness less highly than nontherapy professionals. The two groups did not differ significantly in their preference for warmth.

The relative preferences for the nine concepts expressed by each of the various groups of psychotherapists appear in Table 2. Each concept is listed with its score, ranked according to preference, for each professional group and experience level, and for combined groups of experienced and inexperienced members of each profession. This last comparison involved 96 experienced versus 113 psychotherapists in training.

The total of the *ranks* of the concepts for the six groupings of therapists shows empathy to be clearly most preferred. Dependability and interest are tied for second and are followed by sincerity, respect, warmth, objectivity, sureness, and directiveness, in that order. A slightly different order of preference is produced by summing the *scores* for each concept across all groups, although empathy

TABLE 2
COMPARISON OF CONCEPTS WITHIN SUBJECT GROUPS

Group	Rank								
	1	2	3	4	5	6	7	8	9
Social workers									
Experienced									
Concept	EMP	INT	RES	DEP	SINC	WARM	OBJ	SURE	DIR
Score	24.45	22.15	19.00	18.65	18.50	13.55	11.80	10.15	5.75
Duncan									
Social workers									
Inexperienced									
Concept	EMP	INT	SINC	RES	DEP	WARM	OBJ	SURE	DIR
Score	23.41	20.45	19.50	18.95	18.53	15.02	11.60	8.83	7.88
Duncan									
Psychologists									
Experienced									
Concept	EMP	SINC	DEP	INT	RES	WARM	SURE	OBJ	DIR
Score	23.10	22.00	21.06	20.54	19.30	13.98	10.44	9.44	4.14
Duncan									
Psychologists									
Inexperienced									
Concept	EMP	INT	DEP	SINC	RES	WARM	SURE	OBJ	DIR
Score	25.22	21.15	20.82	20.33	18.96	13.63	9.74	9.11	4.67
Duncan									
Psychiatrists									
Experienced									
Concept	DEP	EMP	SINC	INT	RES	WARM	SURE	OBJ	DIR
Score	22.92	22.39	21.15	20.46	16.58	13.19	11.46	10.50	5.35
Duncan									
Psychiatrists									
Inexperienced									
Concept	EMP	DEP	SINC	INT	RES	OBJ	WARM	SURE	DIR
Score	22.89	20.82	20.29	19.82	15.75	15.54	12.11	10.25	6.54
Duncan									
All social workers									
Concept	EMP	INT	SINC	RES	DEP	WARM	OBJ	SURE	DIR
Score	23.68	20.89	19.24	18.96	18.56	14.64	11.65	9.17	7.33
Duncan									
All psychologists									
Concept	EMP	SINC	DEP	INT	RES	WARM	SURE	OBJ	DIR
Score	23.84	21.42	20.97	20.75	19.18	13.86	10.20	9.33	4.33
Duncan									
All psychiatrists									
Concept	EMP	DEP	SINC	INT	RES	OBJ	WARM	SURE	DIR
Score	22.65	21.83	20.70	20.13	16.15	13.11	12.63	10.83	5.96
Duncan									

Note.—EMP = Empathy, INT = Interest, RES = Respect, DEP = Dependability, SINC = Sincerity, WARM = warmth, OBJ = Objectivity, SURE = Sureness, DIR = Directiveness.

remains the most, and directiveness the least preferred of the concepts. This is presented in the table under "All psychotherapists." The results of Duncan's multiple-range test are shown under the means for each S grouping. Scores joined by a bar indicate preferences

which are *not* significantly different from one another.

Table 3 presents the same means grouped by concepts. Duncan's multiple-range test was used to indicate preferences of the different professional groups which are not significantly

TABLE 3
COMPARISON OF SUBJECT GROUPS WITHIN CONCEPTS

Concept	Rank						<i>F</i>					
	1	2	3	4	5	6	Profession	<i>p</i> <	Experience	<i>p</i> <	Interaction ^a	<i>p</i> <
Objectivity Profession Score Duncan	Pi 15.54	SWe 11.80	SWi 11.60	Pe 10.50	PSYe 9.44	PSYi 9.11	9.77	.005	4.60	.05	6.35	.005
Dependability Profession Score Duncan	Pe 22.92	PSYe 21.06	Pi 20.82	PSYi 20.81	SWe 18.65	SWi 18.53	7.91	.005	1.40	<i>ns</i>	0.85	<i>ns</i>
Sincerity Profession Score Duncan	PSYe 22.00	Pe 21.15	PSYi 20.33	Pi 20.29	SWi 19.50	SWe 18.50	3.69	.05	.55	<i>ns</i>	1.32	<i>ns</i>
Sureness Profession Score Duncan	Pe 11.46	PSYe 10.44	Pi 10.25	SWe 10.15	PSYi 9.74	SWi 8.83	1.34	<i>ns</i>	2.48	<i>ns</i>	0.08	<i>ns</i>
Directiveness Profession Score Duncan	SWi 7.88	Pi 6.54	SWe 5.75	Pe 5.35	PSYi 4.67	PSYe 4.14	4.90	.01	4.05	.05	0.53	<i>ns</i>
Empathy Profession Score Duncan	PSYi 25.22	SWe 24.45	SWi 23.41	PSYe 23.10	Pi 22.89	Pe 22.38	2.56	.10	.80	<i>ns</i>	2.37	.10
Respect Profession Score Duncan	PSYe 19.30	SWe 19.00	PSYi 18.96	SWi 18.95	Pe 16.58	Pi 15.75	14.51	.005	.64	<i>ns</i>	0.20	<i>ns</i>
Interest Profession Score Duncan	SWe 22.15	PSYi 21.15	PSYe 20.54	Pe 20.46	SWi 20.45	Pi 19.82	1.44	<i>ns</i>	1.06	<i>ns</i>	1.42	<i>ns</i>
Warmth Profession Score Duncan	SWi 15.02	PSYe 13.98	PSYi 13.63	SWe 13.55	Pe 13.19	Pi 12.11	1.99	<i>ns</i>	.00	<i>ns</i>	1.22	<i>ns</i>

Note.—Pi = Psychiatrist, inexperienced; Pe = Psychiatrist, experienced; PSYi = Psychologist, inexperienced; PSYe = Psychologist, experienced; SWi = Social worker, inexperienced; SWe = Social worker, experienced.

^a Profession X Experience.

different. *F* values and their significance levels are presented to the right of the concepts.

Focusing on the preferences expressed by psychiatrists, Table 3 indicates that psychiatric residents value objectivity more than all other groups, all psychiatrists value respect less than other groups of therapists, and experienced psychiatrists value sureness more than inexperienced social workers. The com-

parisons for psychologists indicate that inexperienced psychologists value empathy more than do experienced psychologists and all psychiatrists. The comparisons for social workers show that all social workers value dependability and sincerity less than experienced psychiatrists and psychologists, experienced social workers value interest more than inexperienced psychiatrists, inexperienced so-

cial workers value warmth more than inexperienced psychiatrists, and social work students and psychiatric residents both value directiveness more than experienced psychiatrists and all psychologists.

Level of experience does not seem to differentiate relative preferences with the exception of directiveness. Objectivity is partially related to this variable, but the results reflect the interaction with profession; that is, inexperienced psychiatrists' scores seem to determine the greatest portion of this difference.

DISCUSSION

Two central findings of this investigation were that (a) therapeutic qualities of empathy, dependability, interest, sincerity, and respect seem to be most valued by all psychotherapists, and (b) within the range of preferences certain qualities were differentially valued according to the training of the psychotherapists.

The general similarity of the preferences of all groups of therapists is striking. Greater differences might have been anticipated between the subsamples because of variations in kind and amount of training and to a lesser extent in professional self-image and professional goals. Social workers, for example, with their access to the community, and psychiatrists, with their medical training, might both be expected to consider themselves more in a position for active intervention than would psychologists. This may be reflected in the directiveness scores, especially for the inexperienced social workers and psychiatrists. Just as noteworthy, however, is the fact that all therapists placed directiveness at the bottom of their list of preferences. It might also be expected, because of the structured nature of both the analytic approach and setting, that psychiatrists would tend to emphasize technique variables, and social workers, faced with a more fluid and uncontrolled setting, would emphasize relationship variables, with psychologists falling somewhere in between. This was not borne out by the data in any consistent way. Previous studies by Fiedler (1950b) and Strupp (1955a) do in fact emphasize the similarities between professional groupings in conceptualization of an ideal therapeutic relationship or an ideal re-

sponse to a specific therapeutic issue. They differ, however, by demonstrating an experience effect which was not found to be significant here. In addition, Fiedler found that nontherapists were also able to describe an ideal therapeutic relationship while the present data show consistent differences between the values of psychotherapists and non-therapist professionals.

It may be that psychotherapists, because of some self-selection factors, differ in personality structure from members of other professions, and to a lesser degree, from other therapists in related disciplines. The type of training received may have been a major determinant of the response patterns, which in turn may reflect either behavior or merely goals or values. Thus far, from our data, it can only be stated that there are certain general patterns of overall agreement among a broad range of psychotherapists regarding these concepts in the abstract. In addition, there are subpatterns which tend to distinguish therapists with different backgrounds and degrees of experience from one another.

Another possible interpretation of the findings is that an artifact was produced by the context of the test. The consistency with which warmth, objectivity, sureness, and directiveness are found at the bottom of the lists for each group may be taken to imply that they are the more value laden concepts and therefore more subject to negative interpretation. These factors may have been read as "seductiveness, coldness, dogmatism, and authoritarianism" when paired with possibly more moderate or neutral terms. Consequently, such concepts as empathy and dependability were found to be consistently highly rated by the various psychotherapy professions.

Some corroboration regarding these values as conscious preferences among therapists comes from the work of McNair and Lorr (1964). The "directiveness" dimensions in the two studies can be considered roughly equivalent and "warmth" can be treated as a partial reciprocal of their "impersonal" factor. When the data in their table presenting profession by technique pattern are broken down into the single technique factors, social workers are seen to prefer directiveness most often

(71%), with psychiatrists (66%) and psychologists (29%) following in that order. This agrees with our findings for all social workers, psychiatrists, and psychologists. In Table 2 it is shown that the mean score for all social workers on this variable is highest for the three professions, and that all psychologists score lowest. An examination of Table 3 shows that in each case the trainee group is more in favor of directiveness than are the experienced members of their profession. But the order for the experienced and inexperienced groups is the same.

A similar breakdown of their impersonal dimension shows psychologists to be highest on the personal end of the factor (55%), social workers next with 49%, and psychiatrists with 34%. As indicated in Table 2, social workers are highest on warmth, followed by psychologists and psychiatrists. These are data for combined groups, however, and McNair³ and Lorr's sample consisted of experienced therapists. Table 3 indicates that experienced psychologists do in fact choose warmth more often on the PPS than do experienced social workers and psychiatrists.

A number of differences in value preferences among groups of varying professional training or amount of experience have been noted in the results. This is the first step in the direction of understanding what such differences mean, but a variety of related comparisons must be made before meaningful patterns and explanations will emerge. For instance, what are the relationships between the attitudes measured and the behavior of therapists in the practice of psychotherapy? Do different relationships between attitude and practice hold for experienced and inexperienced members of the same profession and of different professions? How much can attitudes and practice be modified by training and experience? How does this vary with the personality of the therapist, the type of attitudes held, and the extremeness of the attitude? Which therapist attitudes and behaviors relate most strongly with positive outcome in psychotherapy and how does this interact with patient diagnosis? There are findings in this area which might be clarified and systematized

by examination in a broader context (Fiedler, 1950a; Strupp, 1955b). The data gathered provide base-line information within an integrated system suitable for such a broad survey of therapists' attitudes and behavior.

The authors have developed a comprehensive set of variables which define a wide range of therapist attitudes and behavior. The particular instrument used in this study may prove of value as a self-rating device, but it is limited to aspects of personality which are nonoffensive or socially desirable. To measure extreme (high or low) features of a therapist's interaction with a patient, other forms should be employed, reflecting the full range of these 10 variables.

To illustrate, one such technique may be used in rating actual therapy behavior from taped interviews or observation of the ongoing therapy. This form would also be suitable to ascertain the ability of supervisors, instructors, and peers to predict or know the therapy behavior of others. Another technique including choices for moderate as well as excessive characteristics may be used to record the patient's perception of his therapist. Lorr (1965) has already demonstrated the potential importance of this technique with the finding that patient perception of therapist understanding and acceptance is significantly related to self-rated improvement and therapist assessment of client satisfaction. Finally, an anticipatory form of the same scale would allow the individual seeking treatment to record his expectations of what the therapist will be like. Heine and Trosman (1960) have already shown the usefulness of patient anticipations in the prediction of dropout from psychotherapy. An outpatient study is currently underway which will attempt to replicate their findings with the expanded list of variables presented here. Conceivably, objective measures of transference and countertransference issues can be obtained by comparing patient anticipations, patient perceptions of therapists, and therapists' self-ratings with the ratings made by objective, trained personnel from actual tape recordings. All scales use the same frame of reference, so that direct comparisons can be made between any or all of the different proposed forms.

³D. M. McNair, personal communication, March, 1968.

REFERENCES

- BANDURA, A. Psychotherapist's anxiety level, self-insight, and psychotherapeutic competence. *Journal of Abnormal and Social Psychology*, 1956, 52, 333-337.
- BERGIN, A. E. Some implications of psychotherapy research for therapeutic practice. *Journal of Abnormal Psychology*, 1966, 71, 235-246.
- EDWARDS, A. L. *Experimental design in psychological research*. New York: Holt, Rinehart & Winston, 1966.
- FIEDLER, F. E. A comparison of therapeutic relationships in psychoanalytic, nondirective, and Adlerian therapy. *Journal of Consulting Psychology*, 1950, 14, 436-445. (a)
- FIEDLER, F. E. The concept of an ideal therapeutic relationship. *Journal of Consulting Psychology*, 1950, 14, 239-245. (b)
- FOX, R. E., & GOLDIN, P. C. The empathic process in psychotherapy: A survey of theory and research. *Journal of Nervous and Mental Disease*, 1964, 138, 323-331.
- HEINE, R. W., & TROSMAN, H. Initial expectations of the doctor-patient interaction as a factor in continuance in psychotherapy. *Psychiatry*, 1968, 23, 275-278.
- KNUEFFER, G., JACKSON, D. D., & KRIEGER, G. Personality differences between more and less competent psychotherapists as a function of criteria of competence. *Journal of Nervous and Mental Disease*, 1959, 129, 375-384.
- LORR, M. Client perceptions of therapists: A study of the therapeutic relation. *Journal of Consulting Psychology*, 1965, 29, 146-149.
- M McNAIR, D. M., & LORR, M. An analysis of professed psychotherapeutic techniques. *Journal of Consulting Psychology*, 1964, 28, 265-271.
- ROGERS, C. R. The interpersonal relationship: The core of guidance. *Harvard Educational Review*, 1962, 32, 416-429.
- STRUFF, H. H. Psychotherapeutic technique, professional affiliation, and experience level. *Journal of Consulting Psychology*, 1955, 19, 97-102. (a)
- STRUFF, H. H. The effect of the psychotherapist's personal analysis upon his techniques. *Journal of Consulting Psychology*, 1955, 19, 197-204. (b)
- STRUFF, H. H., WALLACH, M. S., & WOGAN, M. Psychotherapy experience in retrospect: Questionnaire survey of former patients and their therapists. *Psychological Monographs*, 1964, 78(11, Whole No. 588).
- SUNDLAND, D. M., & BARKER, E. N. The orientation of psychotherapists. *Journal of Consulting Psychology*, 1962, 26, 201-212.
- TRUAX, C. B. Effective ingredients in psychotherapy: An approach to unraveling the patient-therapist interaction. *Journal of Counseling Psychology*, 1963, 10, 256-263.
- TRUAX, C. B., & CARKHUFF, R. R. Experimental manipulation of therapeutic conditions. *Journal of Consulting Psychology*, 1965, 29, 119-124.
- TRUAX, C. B., WARGO, D. G., FRANK, J. D., IMBER, S. D., BATTLE, C. C., HOEHN-SARIC, R., NASH, E. H., & STONE, A. R. Therapist empathy, genuineness, and warmth, and patient therapeutic outcome. *Journal of Consulting Psychology*, 1966, 30, 395-401.
- WHITEHORN, J. C., & BETZ, B. J. Further studies of the doctor as a crucial variable in the outcome of treatment with schizophrenic patients. *American Journal of Psychiatry*, 1960, 117, 215-223.

(Received June 20, 1968)

AVERSION TECHNIQUES IN BEHAVIOR THERAPY:

SOME THEORETICAL AND METATHEORETICAL CONSIDERATIONS

G. TERENCE WILSON¹ AND GERALD C. DAVISON

State University of New York at Stony Brook

Recent psychophysiological research indicates that certain cues may be singularly appropriate as functional conditioned stimuli for certain response systems. Such findings raise into question the trend in the aversion conditioning literature away from chemically produced aversion and toward fear produced by painful electric shock. Although this shift from the chemical to the electrical is being made on sound *functional* grounds, the physiological research on appropriateness of cues suggests that contrary to prevalent assumptions in behavior therapy, one must consider most carefully the *topographical* nature of stimuli and responses. At the metatheoretical level, this research makes still more persuasive the argument that behavior modification as a field be concerned with areas of psychology outside of learning and conditioning.

An important issue in the use of aversion techniques concerns the nature of the noxious stimulus. Earlier versions of aversion treatment favored the use of chemically produced aversion in the form of a nausea-producing drug. (cf. Lemere & Voegtlin, 1950). Current trends, however, almost exclusively emphasize electrical aversive stimulation. Rachman (1965) and Rachman and Teasdale (in press) detail the reasons which have dictated this shift to electrical noxious stimuli. At a practical level, the use of nausea-producing drugs results in an unpleasant and unsavory experience for client and clinical staff alike. At a more theoretical level, Eysenck and Rachman (1965) have suggested that these drugs act as central nervous system neural depressants, and thereby are likely to impede the conditionability of the client in accord with Eysenck's personality theory. Yet these factors are really secondary to the most cogent of Rachman and Teasdale's points, namely, the difficulty in adhering to the *traditional conditioning paradigm* (in both classical and anticipatory avoidance conditioning) when using chemical aversion. The crucial parameters here are temporal, that is, getting optimal interstimulus intervals; sequential, that is, having the CS precede the UCS; the fre-

quency of repetition of CS-UCS pairings; and finally, precision in specifying the intensity and duration of the stimuli. Variable and fluctuating individual differences in reactions to emetic drugs virtually proscribe the precision required by the conditioning paradigm for successful application.

Given their premise of the traditional conditioning paradigm, Rachman and Teasdale's points are both well taken and compelling. However, some recent findings from the physiological psychology literature raise questions for current practices in aversion therapy, and suggest also some metatheoretical considerations for behavior therapy.

Garcia and his associates have recently demonstrated that the procedure of pairing a perceptible cue with an effective reinforcer does not lead automatically to effective associative learning. Rather, it seems that the cue must be "appropriate" for the consequences that ensue. Thus, Garcia and Koelling (1966) have shown that avoidance learning with gastrointestinal disturbances produced by ionizing radiation as the UCS readily transferred to a gustatory stimulus, but not to audiovisual and tactile stimuli. On the other hand, avoidance learning with electric shock as the UCS transferred to the audiovisual and tactile stimuli, but not to the gustatory stimulus. Garcia and Koelling (1967) reported the same differential effect following the injection of a drug. Garcia,

¹Requests for reprints should be sent to G. Terence Wilson, Department of Psychology, State University of New York at Stony Brook, Stony Brook, New York 11790.

McGowan, Ervin, and Koelling (1968) paired either flavor or size of food pellets as conditioned stimuli with either malaise induced by X ray or electric shock as the UCS in four groups of rats. The combination of flavor and illness resulted in a significant conditioned decrease in food consumption, but that of size and illness did not. Conversely, the combination of size and pain produced an inhibition of eating while flavor and pain did not. The authors argue that since flavor is closely related to the chemical composition of food, natural selection would favor associative mechanisms relating flavor and olfaction to the aftereffects of ingestion; and they suggest how effective associative learning depends on central neural convergence of the paired afferent input.

These data appear to discredit the commonly made procedural prescription that, in the treatment of an alcoholic, for instance, shock be used to condition aversion to the sight, smell, and taste of alcohol, to exogenous as well as to endogenous stimuli. It seems that fear responses from a shock UCS may be conditioned only to nongustatory attributes of alcohol in our example, and not to the taste and/or smell of alcohol. Any radical change in these conditioned stimuli, such as could be produced by a different environmental complex, could be expected to lead to the "spontaneous recovery" of the consumption of alcohol (Estes, 1955). The best strategy would seem to be to create a chemically based aversion to the taste and/or smell of alcohol and not to the complex of visual, personal, and other stimuli defining the treatment situation.

Garcia, Ervin, Yorke, and Koelling (1967) and Revusky (1968) have extended this notion of the peculiar "appropriateness" of cues to reinforcers, finding that reinforcement can be delayed well beyond the time interval posited as mandatory for effective reinforcement by the traditional S-R model of associative learning. It will be recalled that one of the major arguments in favor of electrical aversion relates to the temporal control it allows.

Interestingly enough, Lazarus (1968) has independently and from a purely clinical standpoint expressed similar views. He reports

the case of an alcoholic whose alcohol consumption remained recalcitrant to faradic shock but disappeared rapidly when a singularly foul-smelling admixture of smelling salts was substituted as the noxious UCS. He goes on to observe that whereas faradic shock seems "appropriate" when the concern is with visual and/or tactile stimuli, as would be the case in a handwashing compulsion, it may be inappropriate in handling overeating and alcoholic consumption. Of course these and similar clinical observations (e.g., Cautela, 1967; Davison, 1968) are in line with Garcia's psychophysiological findings.

The foregoing raises also some interesting issues on the metatheoretical level. The learning-oriented behavior modification literature, especially the operant, emphasizes the overriding importance of *functional* definitions of problems in distinct contrast to *topographical* (e.g., Ferster, 1965; Staats & Staats, 1963). In the attempt to extrapolate from research with infrahumans to research with humans, learning theory oriented behavior modifiers seem, on reflection, to have had little choice but to stress the functional to the virtual exclusion of the topographical, nay, physiological nature of stimuli and responses. Thus, one terms a response as topographically complex as "walking to the door" an "operant" if it can be shown to relate functionally to antecedent and consequent stimuli in the same fashion as more carefully delineated operants in the Skinner box. The same holds true for behavior therapy approaches which stress the classical conditioning paradigm: Pavlov (1928) himself proposed that "every imaginable phenomenon of the outer world affecting a specific receptive surface of the body may be converted into a CS [p. 88]."

The Garcia research obviously raises the serious question of this preoccupation with functional identities to the exclusion of the specific nature of the particular stimuli and responses within one's presumed conditioning paradigm. This work also illustrates most clearly the possible dangers inherent within a limited conception of "behavior modification" as deriving from "modern learning theory" or, more properly, learning principles. Familiarity with such psychophysiological work as Garcia's (*inter alia*) makes

clear the desirability of expanding the field of behavior modification to include general experimental psychology as a whole (cf. Davison, in press).

REFERENCES

- CAUTELA, J. Covert sensitization. *Psychological Reports*, 1967, 20, 459-468.
- DAVISON, G. C. Elimination of a sadistic fantasy by a client-controlled counterconditioning technique: A case study. *Journal of Abnormal Psychology*, 1968, 73, 84-90.
- DAVISON, G. C. Appraisal of behavior modification techniques with adults in institutional settings. In C. M. Franks (Ed.), *Assessment and status of the behavioral therapies and associated developments*. New York: McGraw-Hill, in press.
- ESTES, W. K. Statistical theory of spontaneous recovery and regression. *Psychological Review*, 1955, 62, 145-154.
- EYSENCK, H. J., & RACHMAN, S. *The causes and cures of neurosis*. London: Routledge & Kegan Paul, 1965.
- FERSTER, C. B. Classification of behavioral pathology. In L. Krasner & L. P. Ullmann (Eds.), *Research in behavior modification*. New York: Holt, 1965.
- GARCIA, J., ERVIN, F. R., YORKE, C. H., & KOELLING, R. A. Conditioning with delayed vitamin injections. *Science*, 1967, 155, 716.
- GARCIA, J., & KOELLING, R. A. Relation of cue to consequence in avoidance learning. *Psychonomic Science*, 1966, 4, 123-124.
- GARCIA, J., & KOELLING, R. A. A comparison of aversions induced by X-rays, toxins, and drugs in the rat. *Radiation Research*, 1967, 7, 439.
- GARCIA, J., MCGOWAN, B. K., ERVIN, F. R., & KOELLING, R. A. Cues: Their relative effectiveness as a function of the reinforcer. *Science*, 1968, 160, 794-795.
- LAZARUS, A. A. Aversion therapy and sensory modalities: Clinical impressions. *Perceptual and Motor Skills*, 1968, 27, 178.
- LEMERE, G., & VOEGTLIN, W. An evaluation of the aversion treatment of alcoholism. *Quarterly Journal for the Study of Alcoholism*, 1950, 11, 199-204.
- PAVLOV, I. P. *Lectures on conditioned reflexes*. New York: International, 1928.
- RACHMAN, S. Aversion therapy: Chemical or electrical? *Behaviour Research and Therapy*, 1965, 2, 289-300.
- RACHMAN, S., & TEASDALE, J. D. Aversion therapy. In C. M. Franks (Ed.), *Assessment and status of the behavioral therapies and associated developments*. New York: McGraw-Hill, in press.
- REVUSKY, S. H. Aversion to sucrose produced by contingent X-irradiation: Temporal and dosage parameters. *Journal of Comparative and Physiological Psychology*, 1968, 65, 17-22.
- STAATS, A. W., & STAATS, C. K. *Complex human behavior*. New York: Holt, 1963.

(Received June 21, 1968)

ENCODING OF ATTITUDE BY A SEATED COMMUNICATOR VIA POSTURE AND POSITION CUES¹

ALBERT MEHRABIAN² AND JOHN T. FRIAR

University of California, Los Angeles

In this study, Ss were requested to imagine themselves in situations involving different kinds of addressees and to sit in the ways in which they would if they were actually interacting with these addressees. The four independent factors in the experiment were encoder-communicator attitude, addressee status, addressee sex, and encoder sex. The dependent measures were eye contact, distance, head orientation, shoulder orientation, leg orientation, arm openness, leg openness, and hand, foot, and trunk relaxation. The findings suggest that the most important variables for the communication of positive attitude are small backward lean of torso, close distance, and more eye contact. Also, there is less sideways lean and more eye contact in communications with high-status as compared to low-status addressees. In addition, for female encoders, less arm openness is an indicator of a higher status of the addressee.

The present investigation is an attempt to explore the salient postural, orientation, and distance cues which have been considered relevant to the communication of attitudes. There have been a number of informal discussions of cues which can be relevant in such communication. Hall (1959, 1963, 1964) delineates several variables which he subsumes under the rubric of proxemics. Some proxemic variables are distance between a communicator and his addressee, degree of directness of orientation of a communicator toward his addressee (i.e., the degree to which a communicator's body is turned in the direction of his addressee versus away from the direction of his addressee), and the presence or absence of touching, or eye contact, between communicator and addressee. Thus, the concept of proxemics subsumes variations in postural and distance variables which relate to the degree of directness or immediacy of interaction between a communicator and his addressee (Mehrabian, 1967). In addition to eye contact and directness of body orientation, Machotka (1965) also denotes acces-

sibility of a communicator's body to the addressee (e.g., openness of arms) as a relevant affect or attitude communicating variable. Directness of body orientation can also be construed as a proxemic variable.

Argyle and Kendon (1967) summarized some of the research literature related to the effects of the above variables in communication. For instance, the importance of eye contact in communication involving positive versus negative attitudes has been demonstrated in several studies. Kendon (1967) suggested that the amount of eye contact which a communicator allots to his addressee can serve the regulative function of allowing varying degrees of emotionality, or expression of attitudes between the communicators. Exline and Winters (1965) found that when Ss were interviewed simultaneously by two Es, they tended to have more eye contact with the E whom they preferred. They also found that Ss have more eye contact with an interviewer when they receive a positive treatment from the interviewer than when they receive a negative treatment which produces feelings of dislike. Argyle and Kendon (1967) report an unpublished study by Weisbrod³ who studied eye contact patterns in a group. She found that a speaker rates those of his addressees who look at him most as being instrumental to his goal and

¹ The present investigation was supported in part by University of California Grant 2189 and in part by United States Public Health Service Grant MH 13509 to the senior author. The authors wish to thank Marty Williams, John Ellickson, Sheldon Ksionzky, and Nancy Burke for their assistance in running the experiment and tabulating the data.

² Requests for reprints should be sent to Albert Mehrabian, Department of Psychology, University of California, Los Angeles, California 90024.

³ R. M. Weisbrod. Looking behavior in a discussion group. Unpublished manuscript, Cornell University, 1965.

as valuing him more than addressees who look at him less. She also found that a speaker feels more powerful when he receives more eye contact from his addressees. The latter finding can also be interpreted to indicate that greater eye contact with a speaker communicates a feeling of higher status to the speaker.

Sommer's (1967) review indicates the relevance of variations in distance and orientation to attitudes of the communicator. Studies by Leipold (1963), Little (1965), Argyle and Dean (1965), and Rosenfeld (1965) consistently provide support for the hypothesis that closer positions are assumed to persons toward whom attitude is positive than those toward whom attitude is negative. Rosenfeld also found that his Ss who were assigned an approval seeking role did not assume a significantly more direct body orientation toward their addressees than Ss who were assigned an approval avoiding role. Steinzor (1950), however, observed that in a discussion group, most of the interaction took place between those who were sitting opposite each other (i.e., direct orientation). There are two possible interpretations of this relationship: Ss tend to sit across from people with whom they like to interact and/or sitting directly opposite someone leads to interaction with that person. The first interpretation is more plausible. Since the individuals in Steinzor's group met for several sessions, the choice of a sitting position in the group could be determined on the basis of attitudes among members of the group. Some of Mehrabian's (1967) findings also provide support for the first interpretation. He found that when eye contact is present, a less direct body orientation of a communicator is interpreted by the addressee as an indication of a less positive attitude toward herself than when there is more direct body orientation. Thus, the existing evidence relating orientation to attitude is equivocal.

Machotka's (1965) findings suggest some relationship between openness of arms and attitude. His Ss rated line drawings displaying different degrees of arm openness of nude figures. Compared to the figures with moderate or very open arm positions, the figures with closed arm positions were judged as

being cold, rejecting, unyielding, shy, and passive.

In addition to the above investigations which relate the proxemic variables to attitudes, there are some studies of the significance of relaxation-tension in the communication of affect, attitude, or status differences, between communicators. Schlosberg (1954), denoted the dimension of sleep-tension as being relevant to the expression of emotions. Ekman (1965) asked Ss to rate photographs of communicators in an interview situation on a tension-relaxation dimension. He found that body cues, relative to head cues, are more suited for the communication of intensity of affect. In the present context, the relationship of body relaxation to attitudes or status relationships is of interest. Goffman (1961) found that in psychiatric staff meetings, tension is greater in lower status participants (interns) than in higher status participants (psychiatrists). The latter finding implies that the degree of body relaxation of a speaker is greater when he communicates to a lower status addressee than when he communicates to a higher status addressee.

Mehrabian (1968) explored the main and interactive effects of several proxemic variables on inferred attitudes. In the experiments, various combinations of posture, orientation, and distance of communicators were presented to Ss who inferred attitudes from these cues. In the photographic stimuli presented to Ss, in addition to postural cues, communicator characteristics such as age or sex were also varied. The findings indicate that when a communicator sits in a relaxed posture, a relatively forward lean of his posture toward his addressee communicates a more positive attitude to him. An open position of his legs and arms communicates a more positive attitude than a closed position of his legs and arms. When a communicator is relaxed and standing, a relatively closer position to his addressee communicates a more positive attitude. A relatively direct orientation to his addressee can communicate either positive or negative attitudes, depending on the age and sex of the communicator and the addressee.

The study to be reported employs an en-

coding methodology. In this study *Ss* were requested to imagine themselves in situations involving different kinds of addressees and to sit in the ways in which they would if they were actually interacting with these addressees. There were four independent factors in the experiment. They were encoder attitude, addressee status, addressee sex, and encoder sex. There were two levels of each of these four factors. The two levels of the attitude factor were communicated to *S* by asking him to imagine an addressee whom he liked versus an addressee whom he disliked. The two levels of the addressee status were communicated to *S* by asking him to imagine an addressee who was of a higher status than himself versus an addressee who was of a lower status than himself. There was a total of 11 dependent measures in the study. These were eye contact, distance, head orientation, shoulder orientation, leg orientation, arm openness, and leg openness. Additional dependent measures relating to relaxation were hand relaxation, foot relaxation, and trunk relaxation as measured by backward and sideways angles of lean.

Drawing on the interpretations of research findings previously reviewed, the hypotheses were that there is more eye contact, smaller distance, and greater relaxation in communications with liked than with disliked addressees, and that there is more eye contact and less relaxation in communications with high-status than with low-status addressees.

METHOD

Subjects

Forty-eight University of California undergraduates served as *Ss*. The *Ss* were equally divided as to sex and served in the experiment as part of introductory psychology course requirements.

Procedure

The procedure employed in the experiment was similar to one reported in Mehrabian (1968). The *S* who was left alone with a set of instructions assumed eight different postures which he considered to be typical of his postures while interacting with Liked-Disliked \times High-Low Status \times Male-Female Addressees. The *S*'s behavior was observed through a one-way mirror by three raters who independently rated the behavior on the 11 dependent measures. The order in which *S* carried out the eight conditions was counterbalanced and was not known to the raters.

RESULTS

Reliability

Three judges rated the behavior of *S* on each of 11 dimensions. Product-moment correlations were obtained between the ratings of each pair of judges for each of the 11 dimensions. Mean correlations over the three pairs of judges for each of the 11 dimensions indicate the following reliability estimates: .81 for eye contact, .99 for distance, .82 for head orientation, .93 for shoulder orientation, .88 for leg orientation, .90 for arm openness, .94 for leg openness, .27 for hand relaxation, .16 for foot relaxation, .85 for backward lean of torso, and .82 for sideways lean of torso. Since the comparisons of the ratings of the three judges indicate significant consistency among their judgments (except in the case of hand and foot relaxation), mean judgments over judges for each of the dependent variables were employed in the analyses to be reported.

Eye Contact

A score of zero was assigned for no eye contact and a score of unity was assigned for the presence of eye contact. A $2 \times 2 \times 2 \times 2 \times 24$ analysis of variance was performed on the mean judgments of eye contact obtained from the three judges. In this design, there were two levels each of attitude, addressee status, addressee sex, and encoder sex, with 24 *Ss* nested under each level of the encoder sex factor and repeated measures on the first three factors.

The results of the analysis of variance of the eye contact scores indicate significant effects due to attitude ($F = 17.4$, $df = 1/46$, $p < .01$), addressee status ($F = 14.8$, $df = 1/46$, $p < .01$), addressee sex ($F = 10.6$, $df = 1/46$, $p < .01$), Attitude \times Addressee Sex ($F = 11.4$, $df = 1/46$, $p < .01$), Addressee Status \times Addressee Sex ($F = 9.5$, $df = 1/46$, $p < .01$), Attitude \times Addressee Status \times Encoder Sex ($F = 10.4$, $df = 1/46$, $p < .01$), and Attitude \times Addressee Sex \times Encoder Sex ($F = 9.4$, $df = 1/46$, $p < .01$).

The main and interaction effects involving the attitude factor will be considered first. The significant effects of attitude, Addressee Sex \times Attitude, and Encoder Sex \times Addressee

Sex \times Attitude are hierarchically related; therefore, a consideration of the Encoder Sex \times Addressee Sex \times Attitude Effect will summarize all the necessary findings related to these significant effects.

Application of the Newman-Keuls technique (Winer, 1962) for comparisons of the eight means involved indicates the following significant differences among the means at the .05 level. For male encoders, the cell means for eye contact indicate that irrespective of addressee sex, there is less eye contact with disliked addressees (.68) than with liked addressees (.86). For female encoders, there is significantly less eye contact with disliked male addressees (.53) than with any of the other three addressee groups (i.e., liked males, .88; liked females, .81; and disliked females, .81). In other words, for female encoders, the degrees of eye contact with liked male, liked female, or disliked female addressees do not differ significantly from each other, while all three differ significantly from the degree of eye contact with disliked male addressees.

The main and interaction effects involving the addressee status factor will next be considered. The significant effects of addressee status, addressee sex, Addressee Status \times Addressee Sex are also hierarchically related; therefore, a consideration of the Addressee Status \times Addressee Sex effect will summarize all the necessary findings relating to these three effects. Application of *t* tests for comparison of the four means involved indicates the following significant differences among the means at the .01 level. Irrespective of encoder sex, there is significantly less eye contact with low-status male addressees (.60) than with any of the other three addressee groups (i.e., low-status females, .79; high-status males, .83; and high-status females, .83).

Finally, the Attitude \times Addressee Status \times Encoder Sex effect needs to be considered. Application of the Newman-Keuls technique for comparisons of the eight means involved indicates the following significant differences among the means at the .05 level. For male encoders, there is more eye contact with liked high-status addressees (.98) than with disliked high-status addressees (.71), liked low-status addressees (.74), or disliked low-

status addressees (.64). The latter three means do not differ significantly among themselves. For female encoders, there is less eye contact with disliked low-status addressees (.54) than with disliked high-status addressees (.80), liked high-status addressees (.83), or liked low-status addressees (.85)—the latter three means not differing significantly among themselves.

Distance

The straight-ahead and lateral distances were used to obtain the total distance of *S* from the addressee. The unit of measurement for distance was the length of a side of the 9 \times 9 inch tiles covering the floor of the experimental room. A 2 \times 2 \times 2 \times 2 \times 24 analysis of variance was performed on the mean judgments of distance obtained from the three judges. The design employed for the present analysis, as well as all subsequent analyses of variance, was identical to that already reported for the analysis of the eye contact measure. The results of the analysis of variance of the distance scores indicate significant effects due to attitude ($F = 49.5$, $df = 1/46$, $MS_e = 5.42$, $p < .01$) and addressee sex ($F = 9.9$, $df = 1/46$, $MS_e = 1.95$, $p < .01$). The mean distance to liked addressees (5.57) is less than the mean distance to disliked addressees (7.24). Also, the mean distance to female addressees (6.18) is less than the mean distance to male addressees (6.63).

Orientation Measures

All orientation measures were obtained in such a way that greater angles correspond to less directness of orientation of the encoder's head, shoulders, or legs toward his addressee. The product-moment correlation of the head orientation scores with the shoulder orientation scores is .90; the correlation of head orientation scores with the leg orientation scores is .86; and the correlation of the shoulder orientation scores with the leg orientation scores is .97. While separate analyses of variance were performed for each of these three dependent measures, due to the high intercorrelations among these measures, only the analysis of variance performed on the

average of head, shoulder, and leg orientation scores will be reported.

Analysis of variance of the average orientation scores shows significant effects due to attitude ($F = 4.2$, $df = 1/46$, $MS_e = 438.2$, $p < .05$), Addressee Sex \times Encoder Sex ($F = 10.8$, $df = 1/46$, $MS_e = 588.4$, $p < .01$), and Attitude \times Addressee Status \times Addressee Sex ($F = 4.9$, $df = 1/46$, $MS_e = 464.9$, $p < .05$). The main and interaction effects involving the attitude factor will be considered first. The significant effects of attitude and Attitude \times Addressee Status \times Addressee Sex are hierarchically related; therefore, a consideration of the triple interaction effect will summarize all the necessary findings relating to these effects. Application of the Newman-Keuls technique for comparisons of the eight means involved indicates only one significant difference at the .05 level. Irrespective of encoder sex, the average orientation away from liked female addressees of low status (39.0°) is greater than the average orientation away from disliked male addressees of high status (24.3°).

In considering the Addressee Sex \times Encoder Sex interaction effect, t tests were applied for comparison of the four means involved and the following significant differences were obtained at the .05 level. The average orientation of male encoders away from the female addressees (31.9°) is greater than the average orientation of male encoders away from male addressees (24.4°). Also, the average orientation of female encoders away from male addressees (40.5°) is greater than the average orientation of female encoders away from female addressees (31.7°). In short, this interaction effect is due to a greater degree of average orientation of an encoder away from an addressee of the opposite sex than that away from an addressee of the same sex.

While, as noted above, the results of the separate analyses of variance performed for each of the head, shoulder, and leg orientation measures are congruent and need not be separately discussed, there is one effect obtained from the separate analysis of variance of the leg orientation measure which is of some interest. In that analysis a significant effect due to the encoder sex factor was obtained ($F = 96.7$, $df = 1/46$, $MS_e = 3,170.9$,

$p < .01$). The leg orientation of female encoders away from their addressees (39.7°) is greater than the leg orientation of male encoders away from their addressees (28.8°).

Openness Measures

All openness measures were obtained in such a way that greater openness corresponded to larger numbers. The product-moment correlation of arm openness scores with the leg openness is $-.02$.

Analysis of variance of the arm openness scores shows significant effects due to addressee status ($F = 5.8$, $df = 1/46$, $MS_e = .73$, $p < .05$) and Addressee Status \times Encoder Sex ($F = 5.8$, $df = 1/46$, $MS_e = .73$, $p < .05$). Since these two effects are hierarchically related, a consideration of the interaction effect will summarize all the necessary findings relating to the effects. Application of t tests for comparison of the four means involved indicates the following significant differences at the .05 level. The mean score for arm openness of female encoders communicating to high-status addressees (1.62) is significantly lower than the mean scores for arm openness in the female encoder and low-status addressee (2.05), male encoder and high-status addressee (1.93), and male encoder and low-status addressee (1.93) conditions. The latter three do not differ significantly from each other.

Analysis of variance of the leg openness scores shows only a significant effect due to encoder sex ($F = 51.8$, $df = 1/46$, $MS_e = 2.87$, $p < .01$). The mean degree of leg openness of male encoders (2.05) is greater than the mean degree of leg openness of female encoders (.80).

Relaxation Measures

These measures were obtained such that greater numbers correspond to greater degrees of relaxation. Since the judgments of hand and foot relaxation is unreliable, only the results for the trunk relaxation indexes (i.e., backward and sideways lean) will be reported. The number of degrees that an S leaned backward from the vertical served as one index. Analysis of variance of these scores shows significance due to attitude ($F = 22.4$, $df =$

1/46, $MS_e = 269.9$, $p < .01$) and encoder sex ($F = 19.6$, $df = 1/46$, $MS_e = 523.7$, $p < .01$). The mean angle of backward lean with liked addressees (1.4°) is less than the mean angle with disliked addressees (9.3°). Furthermore, the mean angle for female encoders (2°) is less than the mean angle for the male encoders (10.5°). These results may be summarized by noting that the torso lean is more backward for disliked addressees than for liked addressees and that the lean of the torso of male encoders is typically further back than the lean of the torso of female encoders.

Analysis of variance of the second trunk relaxation index (i.e., the angle of sideways lean of S's torso away from the median plane of the chair in which he sits) shows significant effects due to addressee status ($F = 26.1$, $df = 1/46$, $MS_e = 65.9$, $p < .01$). Both male and female encoders show greater trunk relaxation in communicating to low-status addressees (8.5°) than in communicating to high-status addressees (4.3°).

DISCUSSION

This study was designed to help determine the salient posture and position variables in the communication of affect to addressees of different status and sex. The study was undertaken with a primarily empirical rather than theoretical orientation and indicated that some of the dependent measures which were employed were significant indexes of attitude or status differences, whereas others were either unreliable or insignificant in the experiment.

Relaxation in the posture of a communicator indicates both his attitude and his relative status to his addressee. An integration of findings obtained in the present study with findings obtained in Mehrabian's (1968) study suggests the following postural indexes for increasing degrees of relaxation. Least relaxation: obvious muscular tension in hands and rigidity of posture as exhibited by a straight back and upright trunk; moderate relaxation: subtle indicators of relaxation such as trunk position, for example, a 20° forward lean and a less than 10° sideways lean, curved back and in the case of females an open arm position; extreme relaxation:

obvious trunk relaxation as indicated by a greater than 20° backward lean and a greater than 10° sideways lean. The findings of the present study, together with the findings of Mehrabian's (1968) study, suggest that there is a curvilinear relationship between the attitude of a communicator toward his addressee and the degree of relaxation manifested by the communicator. Degree of relaxation is either very high or very low for a disliked addressee and is moderate for a liked addressee. It would seem that a very low degree of relaxation, that is, extreme tension, would occur with threatening addressees and a high degree of relaxation would occur with non-threatening addressees who are disliked.

In contrast to its relationship to attitude, relaxation seems to exhibit a linear relationship with status, as follows: Of the three degrees of relaxation (i.e., obvious tension, moderate relaxation, and extreme relaxation) only the latter two degrees are relevant for describing the relationship of relaxation to status of an addressee: there is a high degree of relaxation with a low-status addressee, a moderate degree of relaxation with a high-status addressee, and an intermediate degree of relaxation with peers. These findings are consistent with Goffman's (1961) observations.

In addition to relaxation, distance can also be seen as constituting an indicator of attitude and status relationships with an addressee. The findings of the present study together with findings of earlier studies (Leipold, 1963; Little, 1965; Mehrabian, 1968; Rosenfeld, 1965) suggest that there is an inverse linear relationship between distance and attitude. The relationship of distance to status can be conceptualized with the use of the findings in the present study and with Lott and Sommer's (1967) study which found that distance to addressees who are of either a higher or a lower status than the communicator are greater than distances to peers. In the present study, it was found that distance to high-status, relative to low-status, addressees is not significantly different. It may therefore be hypothesized that distance to an addressee is a curvilinear function of the status of the addressee relative to the self and that it is equally great for low-status and high-

status addressees and relatively small for peers.

Eye contact and directness of orientation toward an addressee can also serve as indexes of attitude and status toward the addressee. The findings are consistent with those of Exline, Gray, and Schuette (1965) and Exline and Winters (1965) and indicate that eye contact generally increases with more positive attitude toward an addressee. They also indicate eye contact is greater for high-status, relative to low-status addressees. As in previous studies (Mehrabian, 1968; Rosenfeld, 1965), directness of orientation toward an addressee is not as clearly an index of attitude and status toward an addressee. It was found that orientation is least direct for female addressees of low status and most direct for disliked male addressees of high status. This finding would suggest that directness of orientation might be partially determined by the threatening quality of the addressee for the communicator and suggests further exploration of this particular interpersonal variable as a determiner of posture and position cues.

The above findings were obtained with a role-playing methodology which requested Ss to assume that they were in specific situations, instead of attempting to create such situations in the experiment. It is interesting to note that whenever findings from other investigations involving different kinds of methods are available, the results of the present study are congruent with those already available. This is reassuring for the status of the remaining findings of the present study for which corroborative evidence is absent.

REFERENCES

- ARGYLE, M., & DEAN, J. Eye-contact, distance and affiliation. *Sociometry*, 1965, 28, 289-304.
- ARGYLE, M., & KENDON, A. The experimental analysis of social performance. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. Vol. 3. New York: Academic Press, 1967.
- EKMAN, P. Differential communication of affect by head and body cues. *Journal of Personality and Social Psychology*, 1965, 2, 726-735.
- EXLINE, R. V., GRAY, D., & SCHUETTE, D. Visual behavior in a dyad as affected by interview content and sex of respondent. *Journal of Personality and Social Psychology*, 1965, 1, 201-209.
- EXLINE, R. V., & WINTERS, L. C. Affective relations and mutual glances in dyads. In S. S. Tomkins & C. Izzard (Eds.), *Affect, cognition and personality*. New York: Springer, 1965.
- GOFFMAN, E. *Encounters*. Indianapolis: Bobbs-Merrill, 1961.
- HALL, E. T. *The silent language*. New York: Doubleday, 1959.
- HALL, E. T. A system for the notation of proxemic behavior. *American Anthropologist*, 1963, 65, 1003-1026.
- HALL, E. T. Silent assumptions in social communication. *Disorders of Communication*, 1964, 42, 41-55.
- KENDON, A. Some functions of gaze-direction in social interaction. *Acta Psychologica*, 1967, 26, 22-63.
- LEIPOLD, W. D. Psychological distance in a dyadic interview. Unpublished doctoral dissertation, University of North Dakota, 1963.
- LITTLE, K. B. Personal space. *Journal of Experimental Social Psychology*, 1965, 1, 237-247.
- LOTT, D. F., & SOMMER, R. Seating arrangements and status. *Journal of Personality and Social Psychology*, 1967, 7, 90-95.
- MACHOTKA, P. Body movement as communication. *Dialogues: Behavioral Science Research*, 1965, 2, 33-66.
- MEHRABIAN, A. Orientation behaviors and nonverbal attitude communication. *Journal of Communication*, 1967, 17, 324-332.
- MEHRABIAN, A. Inference of attitudes from the posture, orientation, and distance of a communicator. *Journal of Consulting and Clinical Psychology*, 1968, 32, 296-308.
- ROSENFELD, H. M. Effect of approval-seeking induction on interpersonal proximity. *Psychological Reports*, 1965, 17, 120-122.
- SCHLOSBERG, H. Three dimensions of emotion. *Psychological Review*, 1954, 61, 81-88.
- SOMMER, R. Small group ecology. *Psychological Bulletin*, 1967, 67, 145-152.
- STEINZOR, B. The spatial factor in face to face discussion groups. *Journal of Abnormal and Social Psychology*, 1950, 45, 552-555.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

(Received June 21, 1968)

MOTIVATIONAL PATTERNS IN THE FAMILIES OF ADJUSTED AND MALADJUSTED BOYS¹

EDWARD J. MURRAY²

University of Miami

ARTHUR SEAGULL

Michigan State University

AND DAVID GEISINGER

Behavioral Therapy Institute

Twenty maladjusted and 20 carefully matched adjusted boys and their parents were administered TATs and interviewed. Objective measures of achievement, affiliation, power, aggression, and experience balance were obtained from the TATs. The two groups of families differed less, motivationally, than expected on the basis of the clinical literature. Paternal aggression and maternal extra-tensiveness in the maladjusted group did suggest some parental conflict. The maladjusted boys perceived their mothers as distant and subtly controlling. There was no evidence of a positive relationship between motivational measures in members of the same family. There was an inverse relation between the maladjusted boy and his parents, particularly the mother.

In the recent focus on the family of maladjusted individuals, it is frequently observed that the needs and conflicts of the parents are of central importance in the development of pathological behavior in the child (Ackerman, 1958; Handel, 1967; Lidz, Fleck, & Cornelison, 1965; Spiegel & Bell, 1959). Thus, for example, a child's difficulty in school may be related to an excessive or a deficient motivation to achieve on the part of one or both parents. Furthermore, it has been suggested that each family has a core problem revolving around some basic motive that is manifest directly or indirectly in each family member. Unfortunately, there is relatively little direct evidence for these phenomena other than clinical impressions.

¹This study was supported by Research Grant M-3209 from the National Institute of Mental Health. The authors wish to express their appreciation to the families who participated in this study as well as to the staffs of the Onondaga County Child Guidance Center, the Syracuse Public School System, the Syracuse Diocese Parochial School System, and the Onondaga County School Systems for their cooperation in obtaining Ss. The authors also wish to thank the following persons for their help in the scoring and statistical analysis: William Claiborn, Ronald Corwin, Jeffrey Elenewski, Susanne Gassner, James Gavin, Roger Greenberg, Bernard Liberman, John McManus, Louisa Murray, Stephen Olins, and Barbara Schiebel.

²Requests for reprints should be sent to Edward J. Murray, Department of Psychology, University of Miami, Coral Gables, Florida 33124.

One of the strategies for studying the needs and conflicts of the family has been the administration of projective tests, particularly the Thematic Apperception Test (TAT), to several members of the same family. Earlier exploratory studies (e.g., Fisher & Mendell, 1956; Rosensweig & Isham, 1947) suggested both direct and inverse relationships between TAT themes in members of the same family. However, there have been few systematic studies using this technique. One attempt was made to relate measures of aggression from TATs administered to members of normal families (Hess & Handel, 1956). A level of aspiration score from the TAT was found useful in differentiating the parents of normal, neurotic, and schizophrenic adults (Fisher, Boyd, Walker, & Sheer, 1959). The conjoint TAT, in which several family members tell stories together, has also been found to differentiate normal and disturbed families (Fisher et al., 1959; Winter, Ferreira, & Olson, 1965, 1966).

The purpose of the present study was to compare the motivational patterns in the families of maladjusted boys with those of a well-matched group of families with adjusted boys. The maladjusted boys were typical of the intact, neurotic-level population frequently referred for treatment by the elementary schools. The TAT was administered individually to the boys and their parents

and then scored for the following motives: achievement, affiliation, power, and aggression. These were selected for two reasons: first, they were typical of the kinds of motives described in the clinical literature on family interaction; and, second, they had the most objectively scored and best validated TAT measures. A TAT measure of experience balance was added because of its relevance to external versus internal sources of motivation. Additional information was obtained through interviews.

It was anticipated that the two groups would be differentiated on the basis of motivational measures for individual family members. It was also expected that there would be a relationship between members of the same family on each of the measures. On the other hand, the level of development of theory in this area was judged as not warranting the statement of directional hypotheses at this time.

METHOD

Subjects

A group of 20 families with a maladjusted boy was obtained through a Child Guidance Center. They were intact, white, lower-middle- or upper-lower-class families, about half Protestant and half Catholic. The fathers were working and the mothers were at home. The parents averaged 12 years of schooling. The boys were between the ages of 8 and 12, of normal intelligence, with problems such as academic failure, anxiety symptoms, rebellious behavior, and social rejection. Boys with mental retardation, neurological involvement, legal delinquency, or psychotic symptoms were eliminated.

A group of 20 matched well-adjusted families were obtained through the school systems. Each adjusted boy was selected from the same grade and school as the maladjusted boy with whom he was matched. The principals eliminated all boys with academic or social adjustment problems. The matching was done on the basis of age, intelligence, social class, parent's education, and religion. Intelligence was estimated on the basis of standard tests in the records. Three graduate students determined social class using the Hollingshead and Redlich Scale (1958) with a 91% agreement. The two groups of families did not differ significantly on any of the matching variables.

Procedure

All Ss were tested and interviewed individually by psychology graduate and senior honors students. The maladjusted families were seen as part of the intake procedure at the Child Guidance Center. The ad-

justed families were contacted by phone and seen at the University.

The TAT was administered with the standard instructions. The cards were presented in the following fixed order: 2, 10, 4, 5, 7BM, New, and 16. The New card showed two boys playing. The stories were tape-recorded, then transcribed and coded for blind scoring.

TAT Measures

Measures were obtained from the coded TAT stories for four motives—achievement, affiliation, power, and aggression. The first three of these are described in detail in Atkinson (1958). The measure of aggression was based on a dissertation by Mac-Casland (1961). The main score is the ratio of direct plus indirect aggression to indirect aggression plus inhibition of aggression.

The four motives were scored by four independent teams of two graduate students each. Each team was carefully trained with the aid of published instructions and practice stories. The reliabilities for the scoring of the four measures with the stories of the families in this study ranged from 86% to 93% agreement. The two groups of families did not differ significantly on length of story.

The stories were also scored for a TAT analog of the traditional Rorschach experience balance developed by Palmer and Lustgarten (1962). The extratensive score is based on descriptions of the external environment and external sources of motivation. The intratensive score is based on expression of internal needs and feelings as well as on well-developed story plots. The final measure used was the ratio of extratensive to intratensive scores. The coded stories were scored by a senior honors student. A reliability check showed 80% agreement.

Interview Measures

Two measures were obtained from six open-ended questions about parents asked of the boys. The first was a rating of the parent-child *relationship* as close or distant. Two judges agreed on 86% of the ratings.

The second measure concerned the *control techniques* used by the parents. "Physically Aggressive" included hitting and spanking, while "Verbally Aggressive" included scolding and yelling. "Nonaggressive" included more subtle control techniques such as restriction and deprivation. Two judges agreed on 87% of the ratings. The data for one child were missing.

RESULTS

TAT Measures

The TAT measures for the fathers in both groups are shown in Table 1. The results were evaluated by two-tailed *t* tests for correlated measures. The measure of aggression clearly differentiates the groups. There are no signifi-

cant differences in achievement, power, or experience balance but there is a trend for a greater affiliation need in the maladjusted fathers ($p < .10$). Thus, the results suggest that fathers of maladjusted boys show a strong aggression motive and, possibly, more affiliation motivation than fathers of adjusted boys.

The TAT measures for the mothers of both the maladjusted and adjusted boys are shown in Table 2. It can be seen that the average scores for the two groups of mothers on affiliation, aggression, achievement, and power are quite similar. None of these differences is significant or even approaches significance. The one measure that clearly differentiates the two groups of mothers is the TAT experience balance. The mothers of maladjusted boys had a high Extratensive-Intrastensive ratio, suggesting that they are relatively more responsive to external than internal cues.

The results for the TAT measures with the neurotic and normal boys are shown in Table 3. The groups do not appear to be significantly different on these measures, with the exception of an almost significant trend for the neurotic boys to show more of an aggression motive ($p < .06$).

Correlational Analysis

A correlational analysis was also made of the TAT measures and is shown in Table 4. Product-moment correlations were computed for the relationship between individuals in each family role, in each of the two groups,

TABLE 2

MEAN TAT MEASURES FOR MOTHERS OF
MALADJUSTED AND ADJUSTED BOYS

Measure	Mothers of mal-adjusted boys	Mothers of adjusted boys	<i>t</i>
Achievement motive	1.85	1.35	.38
Affiliation motive	14.15	13.00	.73
Power motive	5.55	7.22	1.58
Aggression motive	1.82	1.62	.23
Experience balance	.43	.25	3.26*

* $p < .01$, $df = 19$, two-tailed test.

on each of the five measures. An examination of the pattern of correlations for the adjusted families suggests very little relationship between the TAT measures for the fathers, mothers, and boys. At best the pattern suggests a weak positive relationship but only one correlation is significantly different from zero.

An examination of the father-mother correlations in the maladjusted group suggests that there is nothing remarkable about the parental pattern. Except for power motivation, the correlations between the maladjusted parents are not significantly different from zero. Furthermore, the patterns of father-mother correlations in the maladjusted and adjusted groups are quite similar with no significantly different correlations on any of the five measures.

On the other hand, the pattern of correlations between the parents and boys in the maladjusted groups does appear to be different with a strong tendency toward inverse

TABLE 1

MEAN TAT MEASURES FOR FATHERS OF
MALADJUSTED AND ADJUSTED BOYS

Measure	Fathers of mal-adjusted boys	Fathers of adjusted boys	<i>t</i>
Achievement motive	1.00	.95	.03
Affiliation motive	12.05	9.40	1.82
Power motive	6.80	6.05	.88
Aggression motive	1.82	1.40	2.54*
Experience balance	.38	.33	.85

* $p < .05$, $df = 19$, two-tailed test.

TABLE 3

MEAN TAT MEASURES FOR MALADJUSTED
AND ADJUSTED BOYS

Measure	Maladjusted boys	Adjusted boys	<i>t</i>
Achievement motive	-2.10	-.60	1.39
Affiliation motive	7.70	8.10	.50
Power motive	6.35	5.15	1.12
Aggression motive	2.31	1.75	2.05
Experience balance	.20	.17	1.01

TABLE 4

PRODUCT-MOMENT CORRELATIONS OF TAT MEASURES BETWEEN MEMBERS WITHIN THE MALADJUSTED AND WITHIN THE ADJUSTED FAMILIES

Measure	Father-Mother		Father-Son		Mother-Son	
	Maladjusted	Adjusted	Maladjusted	Adjusted	Maladjusted	Adjusted
Achievement motive	.24	-.24	.41	.12	-.10	-.07
Affiliation motive	.16	.18	-.13	.19	-.60**	-.19
Power motive	.48*	.37	-.88**	.00	-.66**	.10
Aggression motive	-.06	-.16	-.18	-.30	-.47*	.22
Experience balance	-.03	.20	.18	.47*	.02	.33

* $p < .05$, $df = 18$, two-tailed test.

** $p < .01$, $df = 18$, two-tailed test.

relationships. The negative correlations for the power motive between maladjusted fathers and boys is substantial and significantly different from zero. The maladjusted father-son power correlation is also significantly different from the corresponding one for adjusted fathers and sons ($z = 4.01$, $p < .001$).

The results are even more striking for the maladjusted mother-son correlations. The correlations for affiliation, power, and aggression show a significant inverse relationship. Two of these are also significantly different from the corresponding correlations in the adjusted group (for power, $z = 2.06$, $p < .01$; for aggression, $z = 2.14$, $p < .05$). Thus, in contrast to the weak positive association in the adjusted group, the motivational measures for the maladjusted boys appear to be inversely related to the parental measures, particularly those of the mother.

The intercorrelations for the five TAT measures for each family role in each of the groups ranged from low negative to medium positive. However, only 5 of these 60 correlations were statistically significant, a result that could be attributed to chance. In addition, the correlations between each TAT measure for a given family role and each different measure for each of the other family roles for each group were also of a chance order. No meaningful patterns within or between groups were noted among the few significant correlations. With correlations this low it is unlikely that a combination of measures would increase the differentiation between the adjusted and maladjusted groups of families.

Interview Measures

The results from the rating of the boys' relationships with their parents as close or distant are shown in Table 5. It can be seen that more maladjusted than adjusted boys describe distant relationships with their parents. The difference between the maladjusted and adjusted boys is significant for the relationship with the mother ($\chi^2 = 5.7$, $df = 1$, $p < .02$). A similar trend can be seen for the boy's perception of the father but the difference between the groups does not quite reach statistical significance ($p < .07$).

The measures for the type of control techniques used by the parents, based on the interviews with the boys, are shown in Table 6. It can be seen that the fathers in the two groups do not differ significantly in type of control used ($\chi^2 = .9$, $df = 3$, ns). On the other hand, the distributions of control techniques used by the maladjusted and adjusted mothers are significantly different ($\chi^2 = 12.1$, $df = 3$, $p < .01$). The adjusted mothers tend

TABLE 5

NUMBER OF MALADJUSTED AND ADJUSTED BOYS DESCRIBING THEIR RELATIONSHIP WITH THEIR PARENTS AS CLOSE OR DISTANT

Description by	Relationship with father		Relationship with mother	
	Close	Distant	Close	Distant
Maladjusted boys	11	8	12	7
Adjusted boys	17	3	19	1

TABLE 6

NUMBER OF MALADJUSTED AND ADJUSTED BOYS DESCRIBING THEIR PARENTS CONTROL TECHNIQUES AS PHYSICALLY AGGRESSIVE, VERBALLY AGGRESSIVE, OR NONAGGRESSIVE

Description of	Description by	Parental control techniques			
		Non-aggressive	Verbally aggressive	Physically aggressive	Insufficient information
Father	Maladjusted boys	5	6	5	3
	Adjusted boys	5	8	3	4
Mother	Maladjusted boys	8	5	3	3
	Adjusted boys	0	14	3	3

to use verbal aggressive techniques, such as scolding and yelling. The maladjusted mothers tend to use the more subtle, nonaggressive techniques of restricting and depriving.

DISCUSSION

Perhaps the most striking result of this study was that the motivational differences between the parents in the two groups were considerably fewer and smaller than anticipated on the basis of the clinical literature. Only paternal aggression and maternal extratensiveness differentiate the groups at a satisfactory statistical level. It is possible that other motivational variables would differentiate the groups. Nevertheless, within the limits of the measures used in this study it would not appear that parents of maladjusted children are grossly different, motivationally, than parents of adjusted children. This supports Frank's (1965) conclusion that when objective measures and adequate controls are used, it is difficult to distinguish between families of emotionally disturbed and normal children.

Although the parents in the two groups are not grossly different, the two significant variables are of some interest. The higher aggression score of the fathers in the maladjusted group is consistent with studies showing greater hostile interaction in maladjusted than adjusted families (Gassner & Murray, 1969; Winter et al., 1966). Since these studies used interaction measures, it was difficult to distinguish the main source of hostility. The

present findings suggest that the more overt hostility is displayed by the father and, to some extent, the boy. The mother may play a more subtle role. The higher extratensive score suggests some difficulty in reacting to internal feelings. In line with this, the maladjusted boys describe their mothers as relatively distant and inhibited in overt aggression. Instead, these mothers seem to use subtle control techniques that could enter into the type of hostile interaction described by Spiegel (1957).

The results from the correlational analysis provide little support for the view that each family has a core problem, at least a core problem captured by these motivational measures. In the adjusted group there is essentially no relationship between family members in regard to motives. The slightly positive trend is similar to the results of Hess and Handel (1956) for a set of TAT aggression measures from normal families. Hess and Handel found the same order of relationship between individuals from the same or different families. In the maladjusted group there is also little relationship between the motives of the parents. In contrast to these findings, the TAT measures for the maladjusted boys tend to be negatively related, substantially in some cases, to the measures for their parents, particularly the mother. Thus, the results provide no support for the idea of direct relationships between motives in the same family and support for the idea of inverse relationships only in the case of maladjusted boys and their parents.

Does any of this help us understand the nature of the difficulties of the maladjusted boys? First of all, it should be noted that while the boys in the two groups behave differently, they do not appear to be very different motivationally. Furthermore, the parents in the two groups are not categorically different. Nevertheless, there does seem to be some hostile tension in the maladjusted families and the boy does seem to be alienated from and reacting against the parents, particularly the mother. Conceivably, the relatively mild psychopathology in the present sample is related to the comparatively small differences in parental variables and the child's reaction to these.

REFERENCES

- ACKERMAN, N. W. *The psychodynamics of family life*. New York: Basic Books, 1958.
- ATKINSON, J. W. (Ed.) *Motives in fantasy, action, and society*. Princeton, N. J.: Van Nostrand, 1958.
- FISHER, S., BOYD, I., WALKER, D., & SHEER, D. Parents of schizophrenics, neurotics, and normals. *Archives of General Psychiatry*, 1959, 1, 149-166.
- FISHER, S., & MENDELL, D. The communication of neurotic patterns over two and three generations. *Psychiatry*, 1956, 19, 44-46.
- FRANK, G. H. The role of the family in the development of psychopathology. *Psychological Bulletin*, 1965, 64, 191-205.
- GASSNER, S., & MURRAY, E. J. Dominance and conflict in the interaction between parents of normal and neurotic children. *Journal of Abnormal Psychology*, 1969, 74, 33-41.
- HANDEL, G. (Ed.) *The psychosocial interior of the family*. Chicago: Aldine, 1967.
- HESS, R. D., & HANDEL, G. Patterns of aggression in parents and their children. *Journal of Genetic Psychology*, 1956, 89, 199-212.
- HOLLINGSHEAD, A. B., & REDLICH, F. C. *Social class and mental illness*. New York: Wiley, 1958.
- LIDZ, T., FLECK, S., & CORNELISON, A. R. *Schizophrenia and the family*. New York: International Universities, 1965.
- MACCASLAND, B. W. The relation of aggressive fantasy to aggressive behavior in children. Unpublished doctoral dissertation, Syracuse University, 1961.
- PALMER, J. O., & LUSTGARTEN, B. J. The prediction of TAT structures as a test of Rorschach's experience-balance. *Journal of Projective Techniques*, 1962, 26, 212-220.
- ROSENZWEIG, S., & ISHAM, A. C. Complementary Thematic Apperception Test patterns in close kin. *American Journal of Orthopsychiatry*, 1947, 17, 129-142.
- SPIEGEL, J. P. The resolution of role conflict within the family. *Psychiatry*, 1957, 20, 1-16.
- SPIEGEL, J. P., & BELL, N. W. The family of the psychiatric patient. In S. Arieti (Ed.), *American handbook of psychiatry*. New York: Basic Books, 1959.
- WINTER, W. D., FERREIRA, A. J., & OLSON, J. L. Story sequence analysis of family TATs. *Journal of Projective Techniques and Personality Assessment*, 1965, 29, 392-397.
- WINTER, W. D., FERREIRA, A. J., & OLSON, J. L. Hostility themes in the family TAT. *Journal of Projective Techniques and Personality Assessment*: 1966, 30, 270-274.

(Received June 24, 1968)

REALITY ATTENTIVENESS-INATTENTIVENESS AND EXTERNALIZATION-INTERNALIZATION IN DEFENSIVE STYLE¹

MARSHALL BUSH,² ROBERT HATCHER, AND MARTIN MAYMAN

University of Michigan

This study attempts to link the degree of attentiveness or inattentiveness one characteristically displays toward his immediate environment and the locus of conflict in subjective experience. It was hypothesized that Ss who avail themselves of opportunities to withdraw attention from reality when such an option is an adaptive alternative in a particular situation would also tend subjectively to experience conflict as lying within themselves, whereas Ss who remain attentionally anchored in their immediate surroundings—even when the task requirements do not demand sustained environmental attentiveness, would tend to experience conflict as lying between themselves and something else. Several measures of different aspects of reality attentiveness-inattentiveness were all found to correlate significantly with an index of externalization-internalization in defensive style among a large sample of female college students.

Several lines of investigation (Guilford & Guilford, 1939; Guilford, 1959; Lacey, 1967; Lacey, Kagan, Lacey, & Moses, 1963; Pribram, 1967; Singer, 1966) point toward the feasibility of conceptualizing a bipolar stimulus-response continuum which is bounded at one end by the cognitive processing of internally produced stimuli and at the other end by the cognitive processing of externally produced stimuli. A convenient way of formulating such a dimension is in terms of a directional parameter of attention deployment along an internal-external stimulus coordinate. In very general terms, internal stimuli refer to such things as memories, images, thoughts, and feelings, while external stimuli refer to the immediate environmental objects of sensation and perception. Voth and Mayman (Mayman & Voth,³ Voth, 1962; Voth &

Mayman, 1963, 1967) have proposed that there are stable individual differences in the distribution of attention cathexes in terms of the priorities assigned to impinging internal and external classes of stimuli. A systematic presentation of this dimension as a cognitive control principle has been attempted elsewhere (Bush, 1968).

Previous research has demonstrated that the study of cognitive control principles can provide access to a more refined understanding of the attentional processes and the cognitive structures underlying such classical defense mechanisms as repression, denial, isolation, and intellectualization (Gardner, 1962, 1964; Gardner, Holzman, Klein, Linton, & Spence, 1959; Gardner & Long, 1962; Holzman & Gardner, 1959; Witkin, 1964). The intent of the present study was to demonstrate the direct relevance of a directional parameter of attention deployment for a basic dimension of defensive style—externalization-internalization. Clinical evidence for such a relationship has been provided by Voth and Mayman (1966, 1967).

Mayman and Voth (see Footnote 3) have coined the term "reality-closeness-reality-distance" in referring to an individual's characteristic degree of attentional embeddedness in his surrounding environment. The reality- to explain inter- and intra-individual differences in receptivity to autokinetic illusion. Unpublished manuscript, 1968.

¹ This paper is based on a dissertation submitted to the graduate school of the University of Michigan in candidacy for the degree of Doctor of Philosophy. The first author wishes to thank his thesis committee, Joseph B. Adelson, Martin Mayman, Frederick Wyatt, Ira Miller, and M. Clemens Johnson, for advice and assistance.

² Requests for reprints should be sent to Marshall Bush who is now at Mount Zion Hospital, Department of Psychiatry, 1600 Divisadero Street, San Francisco, California 94115.

³ M. Mayman and H. M. Voth. Reality-closeness, phantasy, and autokinesis: A dimension of cognitive style. Unpublished manuscript, 1967.

M. Mayman and H. M. Voth. Reality-closeness and autokinesis: Experimental test of an hypothesis

closeness-reality-distance principle describes the directional priorities in attention deployment along an internal-external stimulus dimension in situations where *S* is "free" to attend to "outer" or "inner" sources of stimuli, that is, in situations where the adaptive requirements of the task do not demand active attention to external stimuli or active attention to internal stimuli. The relationship between this approach to individual differences in attentiveness toward the "inner" and "outer" world and other related typologies has been discussed elsewhere (Bush, 1968; Voth & Mayman, 1963).

For purposes of this study externalization-internalization was conceptualized as the *locus of conflict in subjective experience*. An externalizer is an individual for whom conflict experiences are environmentally embedded to an extreme degree, that is, conflict is experienced as lying between the self and something else. An internalizer is an individual for whom the experience of conflict is strongly embedded within the self. This conception of externalization-internalization deals not with particular "mechanisms of defense" but with the phenomenology of conflict (in which a variety of defense mechanisms may be implicated). This formulation of externalization-internalization bears much in common with Horney's (1945) discussion of externalization. The hypothesis to be tested is that reality attentiveness (attentional priority to external stimuli) will be positively related to externalization, while reality inattentiveness (attentional priority to internal stimuli) will be positively related to internalization in defensive style.

The underlying rationale for this prediction is that defensive operations employ the same ego apparatuses and structures that are used in nondefensive cognitive functioning. From one standpoint defenses represent the exploitation of adaptive reality testing processes to reduce awareness of internal sources of threat and distress. Both on a phenomenological and a process description level externalization may be viewed as an outgrowth of a high degree of attentional anchorage in one's immediate external reality. Not only does the attentional priority given to external stimuli sensitize the individual to reality representations of

internal conflict, but the resulting phenomenological world contingent on extreme reality attentiveness is deficient in the categories of subjective experience that are requisite for the apprehension of internal conflict.

Whereas a high degree of continuous reality attentiveness creates a set to refer conflict experiences to environmental circumstances, a high degree of attentiveness to internal thoughts and feelings creates the disposition to recognize and experience conflict as lying within oneself. The individual who habitually gives attentional priority to internal stimuli is better equipped with the experimental structures for cognizing and tolerating conflict within the self. That the externalization-internalization distinction in defensive style is of more than academic interest has been shown by Voth and Mayman's (1966, 1967) findings about the course of psychotherapy with these two different kinds of individuals.

This article reports that portion of a larger investigation (Bush, 1968) which attempted to relate several different aspects of reality attentiveness (rather than several parallel measures of the same variable) to externalization-internalization in defensive style. In particular the distinction was drawn between *response potential* as demonstrated in laboratory tasks and frequency of *behavioral manifestations* of this potential in everyday activities. It was felt that an individual's capacity to withdraw attention from environmental stimuli in an experimental situation would not necessarily show a perfect correspondence with the degree of reality in attentiveness he characteristically exhibits in his daily behavior, although the two should be positively related.

METHOD

Subjects

The *Ss* in this study were 83 female college students at the University of Michigan. The sample included 38 freshmen, 35 sophomores, and 10 juniors. Their ages ranged from 17 to 22 with a modal age of 19. All *Ss* were drawn from the introductory psychology courses and received course credit for their participation.

Measures of Reality Attentiveness and Inattentiveness

The first three measures to be described are based on the autokinetic test and replicate the procedures

employed by Mayman and Voth (see Footnote 3, 1967). For a detailed description of these measures the reader is referred to Mayman and Voth (see Footnote 3, 1967) and Bush (1968).

Extent of perceived autokinetic movement. The rationale for use of the autokinetic test as a measure of reality attentiveness is presented in a series of articles by Voth and Mayman (Mayman & Voth, see Footnote 3; Voth, 1962; Voth & Mayman, 1963, 1966, 1967, 1968). For purposes of the present study, autokinetic responsivity was thought to represent a psychological (and perhaps also neurophysiological) predisposition toward greater ease in the withdrawal of attention from external reality and a heightened sensitivity to inner sources of stimulation. The extent to which such a constitutionally based predisposition will reflect itself in overt behavior is strongly influenced by defensive needs and character style (Voth & Mayman, 1967, 1968).

In the autokinetic test *S* viewed a stationary pin-point of light (.0006 foot candles in intensity) from a distance of 9 feet in a totally darkened room. During the 10-minute test period, *S* was asked to trace with a pencil whatever movement she observed. The *S*'s score was the logarithmic conversion of the length of line drawn to record the apparent motion of the light. (The logarithmic conversion is needed to normalize the distribution of autokinetic scores.) A test-retest reliability coefficient of .83 has previously been reported for this measure of autokinetic perception (Voth & Mayman, 1963).

Loss of awareness of the experimental room. This measure, and the following one to be described, was based on a structured interview conducted immediately following the autokinetic test. Part of the inquiry was intended to ascertain the degree to which *S* preserved a mental image of the experimental room during the 10-minute test period. This procedure was devised by Mayman and Voth (see Footnote 3, 1967) to provide a direct experimental measure of reality-closeness-reality-distance. The degree to which *S* remains aware of the presence of the room, despite its visual obliteration by darkness, is a concrete experimental manifestation of the ease with which he can sustain attention to, or, conversely, withdraw attention from, his immediate surroundings when either alternative is an adaptive option offered by the test situation.

The *S* was asked such questions as, "Can you remember your impressions of the room as you were watching the light?" "To what extent did you remain aware of some of the things in the room?" and "If the room seemed different to you, to what extent did it continue to seem like a room at all?" Responses varied greatly from the extreme of losing all sense of being in any type of enclosure to remaining well aware of the presence and contents of the room throughout the entire 10-minute autokinetic test period.

The protocols were later rated blindly by two of the *Es*⁴ using a 5-point scale. A 1-point rating indicated minimal loss of awareness of the room and a

5-point rating extreme obliviousness to the presence of the room. The index of rater agreement (as measured by the product-moment correlation coefficient) was .83.

Degree of increased subjectivity. This measure was intended to assess the degree to which *S* became immersed in a subjective experience (such as a fantasy or a reminiscence of a past experience) while viewing the autokinetic light. The relevant interview questions inquired about the contents of *S*'s thought processes during the 10-minute autokinetic test period (e.g., "What did the light make you think of?" "When it seemed like you were watching a _____, where did you seem to be watching it from?" and "What other thoughts passed through your mind during the time you were looking at the light?").

The protocols were later rated blindly by one of the *Es* and an advanced graduate student in clinical psychology.⁵ A 5-point scale was employed to indicate the degree to which *S* shifted from a reality-oriented to a highly subjective frame of reference. The 1-point scale position denotes a "realistic" experience of the autokinetic situation while the 5-point scale position indicates that the autokinetic test was transformed into a highly subjective imaginary experience. The correlation between the scores assigned by the two raters was .70.

These two interview-based measures, while focusing on different aspects of attentional withdrawal from one's immediate reality situation, are not completely independent of each other. Overlapping data were used in making each of the ratings. Moreover, a high degree of increased subjectivity necessarily implies some loss of awareness of the experimental room. However, there were enough cases of disparity between the two measures to warrant keeping them separate.

The Reality-Attentiveness-Reality-Inattentiveness (RA-RI) Questionnaire. A 26-item questionnaire (Bush, 1968) was used to measure an individual's characteristic degree of environmental attentiveness. The questions were exclusively restricted to everyday behavioral manifestations of attentiveness or inattentiveness in a variety of situations. The range of behavior covered in the item content included time estimation, spatial orientation, calendrical orientation, experiences of "engrossment," and a variety of memory phenomena which could be attributed to reality attentiveness or inattentiveness. Care was taken to keep all of the items well within the range of normal behavior and to phrase the statements in a way that minimized their social desirability value.

Each item consisted of a statement which *S* was asked to endorse on a 4-point scale (*rarely, somewhat, fairly, or very characteristic of you*). Seventeen items were worded in the direction of reality inattentiveness, 12 items in the direction of reality attentiveness. The following examples typify each kind of item.

⁴ Marshall Bush and Martin Mayman.

⁵ Marshall Bush and Marjorie Freston.

Reality inattentiveness (RI items):

- a. You can become so absorbed in something that you completely lose track of the time.
- b. You momentarily forget what day of the week it is.
- c. Unless you are especially interested in something your mind easily wanders from the topic of conversation.

Reality attentiveness (RA items):

- d. You find it easy to remember the names of people you have just been introduced to.
- e. You know almost exactly how much money you have with you at all times.
- f. You are a naturally good score keeper at games and other activities.

Separate scores were obtained for those items worded in the direction of reality attentiveness and for those items worded in the direction of reality inattentiveness. A composite total score was obtained by summing these two-part scores. Thus each *S* had an RA score, an RI, and a total RA-RI score. In calculating the total score, the scoring was reversed for the RA questions so that a high total score indicates reality inattentiveness and a low total score reality attentiveness. The odd-even split-half correlation coefficient for this questionnaire was .80. Application of the Spearman-Brown formula produced a reliability coefficient of .89.

Measure of Externalization-Internalization

The Conflict Situations test. In order to assess externalization-internalization in defensive style a projective technique was devised, the Conflict Situations test. This test was presented in the form of an opinion survey wherein *Ss* were asked to illustrate through written stories what they felt were the three most difficult crises or personal problems faced by girls during their college years. The amount of time allotted for this test, 30 minutes, proved to be insufficient given the nature of the task. Several *Ss* could not complete three stories in the time allowed. Consequently only the first story is used in the data analysis.

The first story written by each *S* was given two ratings. The first rating was based on a 5-point scale indicating the degree to which the experience of conflict is portrayed as occurring within the self, as opposed to being primarily situational, that is, embedded in particular relationships, situations, or events. The second rating, again using a 5-point scale, was based on how attention is distributed within the story between the internal and external referents of the conflict situation. A high rating here indicates that a large proportion of the story was devoted to describing the subjective side of the conflict experience in the central character. The *S's* score was simply the sum of these two ratings. A high score represented internalization, and a low score externalization in defensive style.

Both ratings were made for every *S's* first story by two experienced clinical psychologists.⁶ The mean of the externalization-internalization scores assigned by the two raters served as the measure of defensive style for each *S*. Due to several cases in which the raters showed wide disagreement, the correlation between the scores assigned by the two raters was only .57. To evaluate the effect of eliminating those *Ss* whose stories produced the largest rater disagreement, a separate analysis was performed after excluding all cases in which the raters showed a disagreement of four or more points in the assignment of externalization-internalization scores. The net result was to eliminate nine *Ss* (leaving an *N* of 74) while increasing the reliability of the measure of defensive style for the rest of the sample from .57 to .71.

Procedure

The *Ss* were seen for two experimental sessions. In the first 1-hour testing session, *Ss* were individually administered the autokinetic test, an autokinetic test interview, and the Conflict Situations test. The RA-RI Questionnaire was administered during a second testing session. The participants had no prior knowledge about the purpose of the experiment nor of the fact that the two testing sessions were part of the same study. A few *Ss* did not complete their first testing session, while several others were unable to attend the second session, with the result that only partial data were available in nine cases.

RESULTS

Table 1 presents the correlations between the measures of reality attentiveness-inattentiveness and the measure of defensive style. Despite the rather low reliability of the externalization-internalization score, it nonetheless correlates with four of the measures of attentiveness-inattentiveness at the .05 level of significance, and with the remaining two measures at the .01 level of significance (two-tailed probability). When *Ss* producing the largest rater disagreement are eliminated from the samples (which increases the reliability of the measure of defensive style from .57 to .71) a slight increase occurs in the magnitude of the correlations between the measures of attentiveness-inattentiveness and the externalization-internalization score. In all cases the correlations in Table 1 are in the predicted direction, confirming the predicted relationship between attentional anchorage in one's immediate environment and the tendency to externalize the experience of conflict.

While it was not possible to directly control

⁶ Kay Tooley and Naomi Lohr.

for the effects of intelligence on the observed results, verbal ability scores and cumulative grade point averages were obtained for all Ss in the samples. The verbal ability score from the Scholastic Aptitude Test correlated significantly with the RA subscore and the total score on the RA-RI Questionnaire ($-.36$ and $.24$, respectively), but did not correlate significantly with the measure of defensive style or the other measures of reality attentiveness-inattentiveness. Cumulative grade point average correlated significantly ($-.29$) with only one variable, degree of increased subjectivity. Neither verbal ability, nor academic achievement seems to play a significant role in the obtained relationships between attention deployment and defensive style.

Table 2 presents the intercorrelations between the several measures of reality attentiveness-inattentiveness. Disregarding for the moment the correlations involving the autokinetic score, the other measures of reality attentiveness-inattentiveness (with one exception) all show significant intercorrelations in the predicted direction. The $.59$ correlation between the mean rating for loss of awareness of the room and the mean rating for

degree of increased subjectivity indicates that these two interview-based measures are closely related to each other; the withdrawal of attention from the surrounding environment appears to be accompanied by an internally directed attentiveness to one's own thoughts, fantasies, memories, and feelings.

The correlations between the mean rating for loss of awareness of the room and the questionnaire measures of habitual reality attentiveness-inattentiveness ($.34$, $.30$, $-.30$) indicate that a person's attentional priorities in a laboratory task are reflective of his characteristic behavior in everyday life. This relationship is, however, a modest one. Other factors would seem to mediate the degree to which the response potential evidenced in an experimental situation becomes a prominent feature of one's everyday behavior. The mean rating for degree of increased subjectivity correlates significantly with only the RA items on the RA-RI Questionnaire, although the correlation with the RI items is in the expected direction. The $-.47$ correlation between the RI and RA items is only moderate, suggesting that the "reality inattentive" questions and the "reality attentive" questions, while tapping an overlapping domain, still possess some independent sources of variance not shared by the other.

One unexpected finding in Table 2 was the lack of significant correlations between the autokinetic score and the other measures of reality attentiveness-inattentiveness. No satisfactory explanation for this result is available. Further analysis (Bush, 1968) showed that the relationship between the amount of perceived autokinetic motion and the RA-RI Questionnaire was markedly curvilinear. The Ss who perceived the least amount of movement (2-35 inches), as well as those who perceived the most movement (71-400 inches), received high scores for reality attentiveness and low scores for reality inattentiveness, while the middle group (36-70 inches) received low scores for reality attentiveness and high scores for reality inattentiveness. To put the matter another way, a high score in the direction of reality inattentiveness on the RA-RI Questionnaire was associated with a moderate to high degree of autokinetic perception, whereas a high score in the direction

TABLE 1

INTERCORRELATIONS BETWEEN MEASURES OF REALITY
ATTENTIVENESS-INATTENTIVENESS AND
EXTERNALIZATION-INTERNALIZATION

Measure	Mean externalization-in-internalization score using all data on all Ss ^a	Mean externalization-in-internalization score with Ss producing largest rater disagreement removed ^b
Log autokinetic tracing	.26*	.32**
Loss of awareness of room (mean rating)	.29*	.31**
Degree of increased subjectivity (mean rating)	.28*	.33*
RA-RI Questionnaire total score	.34**	.38**
RI subscore	.31**	.36**
RA subscore	-.27*	-.27*

^a *N* varies between 77 and 83 due to missing data.

^b *N* varies between 69 and 72 due to missing data.

* $p < .05$, two-tailed.

** $p < .01$, two-tailed.

of reality attentiveness could be associated either with a low degree or with an extremely high degree of autokinetic perception.

Some preliminary efforts to control for the effects of individual differences in drawing style on autokinetic tracings yielded a most unusual finding. A drawing style factor appeared to be operative only among Ss recording the greatest amounts of autokinetic movement. Among the 19 Ss who recorded more than 70 inches of movement, the correlation between the index of drawing style and the autokinetic score was .46 (.05 level of significance). For the remainder of the sample, the correlation between drawing style and autokinesis was .05. The effects of eliminating from the sample the 19 Ss recording the highest degree of movement are presented in Table 3. While the relationships between the autokinetic test, loss of awareness of the room, and degree of increased subjectivity remain essentially unchanged, the relationship between the autokinetic scores and the RC-RD questionnaire is made linear. The autokinetic test now correlates significantly with the total RA-RI Questionnaire score and the RI subscore, and there is an increase in the correlation between the RA-RI Questionnaire and

TABLE 2

INTERCORRELATIONS AMONG MEASURES OF REALITY
ATTENTIVENESS-INATTENTIVENESS USING ALL
AVAILABLE DATA ON ALL SUBJECTS

Measure	2	3	4	5	6
1. Log autokinetic tracing	.17	.18	.07	.10	.01
2. Loss of awareness of the room (mean rating)		.59***	.34**	.30**	-.30**
3. Degree of increased subjectivity (mean rating)			.27*	.20	-.31**
4. RA-RI Questionnaire total score				.94***	-.75***
5. RI subscore					-.47***
6. RA subscore					

Note.—N varies between 75 and 83 due to missing data.

* $p < .05$, two-tailed.

** $p < .01$, two-tailed.

*** $p < .001$, two-tailed.

TABLE 3

INTERCORRELATIONS AMONG MEASURES OF REALITY
ATTENTIVENESS-INATTENTIVENESS AFTER SUBJECTS
WITH AUTOKINETIC SCORES ABOUT 70 INCHES
HAVE BEEN ELIMINATED

Measure	2	3	4	5	6
1. Log autokinetic tracing	.18	.14	.32*	.30*	-.23
2. Loss of awareness of the room (mean rating)		.59***	.43***	.41***	-.33*
3. Degree of increased subjectivity (mean rating)			.28*	.21	-.33*
4. RA-RI Questionnaire total score				.94***	-.76***
5. RI subscore					-.50**
6. RA subscore					

Note.—N varies between 58 and 63 due to missing data.

* $p < .05$, two-tailed.

** $p < .01$, two-tailed.

*** $p < .001$, two-tailed.

loss of awareness of the room. One implication of these results which needs to be pursued in future research is that the degree and type of relationships exhibited by the autokinetic test, as well as the influence of a drawing style factor, may vary according to the range of scores within which one is working.

DISCUSSION

The results, in general, suggest that a relationship does exist between the attentional priorities one assigns to internally and externally produced stimuli and the locus of conflict in subjective experience. The Ss who readily withdraw attention from their immediate environment and attend instead to internal states—when such behavior is an adaptive option afforded by the response demands in a particular situation—also subjectively experience conflict as something which is embedded within themselves. Conversely, Ss who sustain attention to their external surroundings when such behavior is not necessitated by the adaptive requirements in a particular situation, also subjectively experience conflict as lying between themselves and something else. In describing conflict they focus on the

external trappings of the situation, rather than on the internal sources of tension. The externalization-internalization dimension appears to be of decisive importance in determining an individual's capacity to respond to analytically oriented "insight" therapy (Voth & Mayman, 1966, 1967). Externalizers view the source of their problems as stemming from without and find it very difficult to think psychologically and make effective use of insight.

It is noteworthy that all of the measures of reality attentiveness-inattentiveness correlated to a comparable degree with the externalization-internalization score, despite the fact that they correlated only very modestly with each other. It seems likely that we are dealing here with different but related aspects of attentional priorities for internal and external sources of stimulation. The nonsignificant relationships between the autokinetic test and the other measures of attention deployment is very puzzling in light of previous experimental results (Mayman & Voth, see Footnote 3) and not readily explainable. One factor which seems to have played an important part in this finding was the fact that this sample, in contrast to previous ones, lacked Ss who perceived no autokinetic movement. This restriction on the range of autokinetic scores at the "low-movement" end of the continuum may well have prevented the expected relationships from appearing. In a more heterogeneous sample the relationships which were observed might be found to be stronger than appeared in this study. Another unknown stemming from the use of a very homogeneous sample of female college students is the possible effect of sex differences on the variables under investigation. Clarification of this issue must await further study.

We are dealing here with two domains of behavior which have been sampled in only a preliminary and incomplete manner. The locus of conflict in subjective experience is only one aspect of externalization-internalization in defensive style. Other related aspects of coping with conflicts include how responsibility for personal difficulties gets distributed, the kind of symptomatology one is likely to develop, the kind of resolution which is sought after, and expectancies concerning the control

of reinforcement (Rotter's, 1966, internal-external control dimension). Similarly, with regard to the kind and degree of attentiveness one displays toward his inner states and outer environment, we have dealt with only a limited segment of a very complex area of human behavior. People differ in the areas of reality toward which they remain attentive and in the modalities through which they attend, just as they differ in the areas of self-experience toward which they are characteristically attentive or inattentive. As the dimensionality of reality attentiveness-inattentiveness and externalization-internalization in defensive style is further explored and understood, it should be possible to link particular components of attentiveness-inattentiveness with specific facets of externalization-internalization. A more differentiated understanding of the enduring parameters of attention deployment will not only shed light on the defensive operations (and other ego structures and functions) which use these attentional processes, but may in addition provide the key to a more adequate understanding of the individual's phenomenological world.

REFERENCES

- BUSH, M. A study of reality-closeness-reality-distance: A directional determinant of attention deployment. Unpublished doctoral dissertation, University of Michigan, 1968.
- GARDNER, R. W. Cognitive controls in adaptation: Research and measurement. In S. J. Messick & J. Ross (Eds.), *Measurement in personality and cognition*. New York: Wiley, 1962.
- GARDNER, R. W. The development of cognitive structures. In C. Scheerer (Ed.), *Cognition*. New York: Harper & Row, 1964.
- GARDNER, R. W., & LONG, R. I. Cognitive controls of attention and inhibition: A study of individual consistencies. *British Journal of Psychology*, 1962, 53, 381-388.
- GARDNER, R. W., HOLZMAN, P. S., KLEIN, G. S., LINTON, H. B., & SPENCE, D. P. Cognitive control. *Psychological Issues*, 1959, 1(4, Whole No. 4).
- GUILFORD, J. P. *Personality*. New York: McGraw-Hill, 1959.
- GUILFORD, J. P., & GUILFORD, R. B. Personality Factors D, R, T, and A. *Journal of Abnormal and Social Psychology*, 1939, 34, 21-36.
- HOLZMAN, P. S., & GARDNER, R. W. Leveling and repression. *Journal of Abnormal and Social Psychology*, 1959, 59, 151-155.
- HORNBY, K. *Our inner conflicts*. New York: Norton, 1945.

- LACEY, J. I. Somatic response patterning and stress: Some revisions of activation theory. In M. H. Appley & R. Trumbull (Eds.), *Psychological stress*. New York: Appleton-Century-Crofts, 1967.
- LACEY, J. I., KAGAN, J., LACEY, B. C., & MOSS, H. A. The visceral level: Situational determinants and behavioral correlates of autonomic response patterns. In P. H. Knapp (Ed.), *Expressions of the emotions in man*. New York: International Universities Press, 1963.
- PRIBRAM, K. H. The new neurology and the biology of emotion: A structural approach. *American Psychologist*, 1967, 22, 830-838.
- ROTTER, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 1966, 80(1, Whole No. 609).
- SINGER, J. L. *Daydreaming*. New York: Random House, 1966.
- VOTH, H. M. Choice of illness. *Archives of General Psychiatry*, 1962, 6, 149-156.
- VOTH, H. M., & MAYMAN, M. A dimension of personality organization. *Archives of General Psychiatry*, 1963, 8, 366-380.
- VOTH, H. M., & MAYMAN, M. The psychotherapy process and its relation to ego-closeness-ego-distance: Part I. *Journal of Nervous and Mental Disease*, 1966, 143, 324-337.
- VOTH, H. M., & MAYMAN, M. Diagnostic and treatment implications of ego-closeness-ego-distance: Autokinesis as a diagnostic instrument. *Comprehensive Psychiatry*, 1967, 8, 203-216.
- VOTH, H. M., MAYMAN, M., CANERO, R., & COYNE, L. Autokinesis and character style: A clinical-experimental study. *Journal of Psychiatric Research*, 1968, 6, 51-66.
- WITKIN, H. A. Origins of cognitive style. In C. Scheerer (Ed.), *Cognition*. New York: Harper & Row, 1964.

(Received June 24, 1968)

Manuscripts Accepted for Publication in the
Journal of Consulting and Clinical Psychology

- Frequency of Dream Recall Estimated by Three Methods and Related to Defense Preference and Anxiety: David Benjamin Cohen*: 408 E. Kingsley, Ann Arbor, Michigan 48104.
- Follow-Up of 812 Girls 10 Years after High School Dropout: Starke R. Hathaway,* Phyllis C. Reynolds, and Elio D. Monachesi: Medical School, Division of Clinical Psychology, University of Minnesota, Minneapolis, Minnesota 55455.
- Self-Care Unit: Some Lessons in Institutional Power: Alan P. Towbin*: The Learning Center, School Street, Farmington, Connecticut 06032.
- Relationship of Academic Underachievement to Incidental Learning: Michele S. Mondani* and Thomas A. Tutko: San Jose State College, 125 South Seventh Street, San Jose, California 95114.
- Personality Characteristics of Parents of Neurotic, Aggressive, and Normal Preadolescent Boys: Lorna M. Anderson*: 330 Prospect Avenue South, Minneapolis, Minnesota 55419.
- Sex Differences and Variations in Body Attitudes: Richard M. Kurtz*: Department of Psychology, Washington University, St. Louis, Missouri 63130.
- Relationship of Work Therapy to Psychiatric Length of Stay and Readmission: Margaret S. Barbee,* K. L. Berry, and Leo A. Mick: Fort Logan Mental Health Center, 3520 West Oxford Center, Denver, Colorado 80236.
- Identification of Clinically Deviant Dimensions of Children's Behavior: Jacob O. Sines,* Jerome D. Pauker, Lloyd K. Sines, and David R. Owen: Department of Psychiatry, University of Missouri, Columbia, Missouri 65201.
- Distinctive Personality Attributes of Child-Abusing Mothers: Barry Melnick and John R. Hurley*: Department of Psychology, Michigan State University, East Lansing, Michigan 48823.
- Reappraisal of the Validity of WAIS, WISC, and WPPSI Short Forms: A. B. Silverstein*: Pacific State Hospital, Box 100, Pomona, California 91766.
- Comparison of Indicators of Homosexuality on the MMPI: Michael I. Singer*: Undercliff Mental Health Center, Meriden, Connecticut 06450.
- Peabody Picture Vocabulary Test as a Measure of General Adult Intelligence: Donald A. Pool* and Robert Brown: Psychology and Vocational Rehabilitation, University of Texas Medical School at Dallas, Dallas, Texas 75235.
- Relationship of Manifest Anxiety to Repression-Sensitization on the MMPI: Patrick F. Sullivan* and Lynn K. Roberts: Department of Psychology, Albion College, Albion, Michigan 49224.
- Family Size and Birth Order as Determinants of Scholastic Aptitude and Achievement in a Sample of Eighth-Graders: Mark Oberlander,* Noel Jenkin, Kevin Houlihan, and John Jackson: Institute for Juvenile Research, 232 East Ohio Street, Chicago, Illinois 60611.

* Asterisk indicates author for whom address is supplied.

VALIDATIONAL STUDY OF MARITAL STATUS AND THE SELF-REPORT SCALE FOR PROCESS- REACTIVE SCHIZOPHRENIA

DONALD H. MEICHENBAUM¹

University of Waterloo, Ontario

Evidence for both predictive and construct validity for the self-report scale and marital status was obtained for 62 male schizophrenic patients. Six-month, 1-year, and 2-year follow-ups indicated that the self-report scale and marital status were equally efficient in predicting length of hospitalization. The inter-correlations between the criterion measures indicate that length of present hospitalization is the best predictor of length and future hospitalization. A consistent finding for both male schizophrenic and male normals was a significant relationship between level of cognitive ability and marital status. Married and divorced males, both schizophrenic and normal, performed on a superior level on the proverbs test and gave more common associations than single males. Possible factors contributing to the efficacy of marital status in predicting prognosis were discussed.

The process-reactive continuum has received much attention in reducing heterogeneity and predicting prognosis of patients diagnosed schizophrenic (Garmezy, 1965; Herron, 1962). The measurement of this dimension has been based on either the Elgin (Becker, 1956; Wittman, 1941) or Phillips (1953) scales. Both scales are time consuming, require a great deal of effort, and are subject to extraneous sources of variance, for example, the varied quality of social and medical records. These limitations were overcome by Ullmann and Giovannini (1964) who developed a 24-item, true-false self-report scale which obtains biographical information to reflect the nature of the patient's prior social interaction with his environment. The procedure for selecting items was based on face validity as judged by three clinicians and by internal consistency criteria. No external criterion was used either in the scale's development or to demonstrate its predictive validity. Construct validity for the self-report scale was provided by Johnson and Reis (1967) and Held and Cromwell (1968) who found a significant correlation between the self-report scale and Phillip's scale and by Ober (1966) who found a significant correlation between the self-report scale and the

Zigler and Phillips (1962) Social Competence Scale.

The purpose of the present investigation was to examine the predictive and construct validity of the process-reactive self-report scale. A second purpose was to compare the predictive effectiveness of the self-report scale with marital status.

METHOD

Subjects

Sixty-two male schizophrenics from the Veterans Administration Hospital, Danville, Illinois, who had been previously involved in other research (Meichenbaum, 1966a) were selected from acute admission wards and back wards. The Ss represent a stratified sample, in terms of age, level of education, and marital status, of patients diagnosed schizophrenic at the hospital. The Ss whose diagnosis of schizophrenia may have been questionable were not included. At the time when the self-report scale was administered the average age of sample was 38.40 with a range of 24-48; average length of present hospitalization was 12.50 months with a range of 1-48 months; mean number of hospitalizations was one and one-half; and average level of education was 11.96 years with a range of 6-16 years. The present sample represents a broad range of pathology including a number of extreme process and reactive schizophrenics.

Procedure

Administration of self-report scale. Pilot research indicated that schizophrenic patients often misinterpreted the questions, or could not support the answer

¹Requests for reprints should be sent to Donald Meichenbaum, Department of Psychology, University of Waterloo, Waterloo, Canada.

they offered. Thus, the self-report scale² was administered verbally to each S. This procedure was followed to insure the likelihood of accurate reporting and to expand the range of Ss to whom the self-report scale could be administered. Belligerent and disoriented patients who were unwilling or unable to take the test on their own were testable in this manner.

Marital status. One of the major classes of items in many prognostic scales such as the Phillips (1953), Modified Elgin Scale (Becker, 1956; Wittman, 1941), and the self-report scale (Ullmann & Giovannini, 1964) concerns the heterosexual behavior and marital status of the patient. Several investigators (Farina, Garnezy, & Barry, 1963; Garfield & Sundland, 1966) have reported that marital status has proved a successful prognostic indicator. Thus, the relative effectiveness of marital status in predicting outcome was examined. The relationship between a 3-point marital scale (1-married, 2-divorced, 3-single) and the criterion was examined. In the present sample 18 Ss were married, 10 were divorced, and 34 were single.

Criterion of predictive validity. The criterion of predictive validity employed was the total number of days the patient spent in a neuropsychiatric hospital (TDH) from the day the self-report scale was administered until a 6-month, 1-year, and 2-year follow-up was obtained. Social service and medical files were examined to keep track of hospital discharge and possible rehospitalizations. The discharge of the patient from the hospital was entirely independent of his score on the self-report scale. A positive correlation between process-reactive scores and TDH means that those Ss who scored toward process end remained in the hospital longer. A positive correlation between marital status and TDH means that those Ss who are single remain in the hospital longer.

Assessment of construct validity. One kind of evidence for the construct validity of prognostic indicators of the process-reactive continuum is a significant correlation with level of thinking pathology, namely, those Ss who indicate a more adequate level of prior social adjustment do better on conceptual tasks; whereas those Ss who evidence less adequacy do more poorly on conceptual tasks. In order to assess this relationship, three cognitive measures were administered: the Vocabulary subtest of the 1960 Stanford-Binet, Form L-M; the Similarities Subtest of the WAIS; and Kaufman Proverbs Test (1960), Form I. The size of the sample for the assessment of construct validity was expanded to include 48 additional schizophrenic Ss from the same hospital

who were used in another research project (Meichenbaum, 1966b). The additional Ss were more chronic and manifested a greater degree of schizophrenicity, especially pathognomic verbalizations. The inclusion of these additional Ss provides a broader range of pathology to determine the relationship between cognitive functioning and process-reactive scores. The relationship between cognitive functioning and marital status was also examined.

RESULTS

Table 1 presents the intercorrelations between process-reactive self-report scores, marital status, and the criterion total number of days hospitalized (TDH). Four TDH scores are reported, namely, the number of days hospitalized in first 180 days; the number of days hospitalized in first 365 days; the number of days hospitalized in the second 365 days; and the total number of days hospitalized over the 2-year period. The set of intercorrelations helps to answer three questions: (a) What is the relative effectiveness of the process-reactive self-report scale and marital status in predicting TDH? (b) What is the relationship between the self-report scale and marital status? (c) What is the relationship between length of present hospitalization and length of future hospitalizations?

The significant correlations between self-report scale and TDH, and marital status and TDH ($p < .01$, $df = 60$), indicate that both the self-report scale and marital status predict the length of hospitalization of patients diagnosed schizophrenic. The differences in the correlations between process-reactive scores and the criteria, and marital status and the criteria, were not significant, although the self-report scale had a consistent trend of higher relationships with the criteria.

TABLE 1

INTERCORRELATIONS OF PROCESS-REACTIVE
SCORES, MARITAL STATUS, AND TOTAL
DAYS HOSPITALIZED

Correlates	Marital status	TDH 1-180	TDH 1-365	TDH 366-730	TDH 1-730
Process-reactive scale	.776	.411	.441	.341	.393
Marital status		.325	.389	.323	.369
TDH 1-180			.904	.711	.837
TDH 1-365				.817	.944
TDH 366-730					.961

Note.— $N = 62$; all correlations significant, $p < .005$, one-tailed.

² Question 15 was changed to read "More than once in the last year after some group meeting I have talked with some other members about something that went on." The phrase "I have stayed on" was deleted because the patients in group meetings at the Veterans Administration Hospital are typically ushered out of the meeting room at the completion of the group session.

The high-positive relationship ($r = .776$) between process-reactive scores and marital status is supported by the mean process-reactive scores for married ($N = 18$), divorced ($N = 10$), and single ($N = 34$) patients, respectively, 6.11, 10.40, and 15.62, all significantly different from each other ($p < .01$). Four of the 24 items on the self-report scale concerns marital status and several items are highly correlated with marital status, contributing to the high correlation between the two measures.

The high intercorrelations between the respective TDH measures indicate that length of present hospitalization is a good predictor of length of future hospitalization. The correlation of .904 between TDH 1-180 and TDH 1-365 is somewhat inflated because of the redundancy of information in the latter measure. However, when one partials out the first 180 days from the first year score, the correlation between first 6 months and second 6 months drops only to .820, still a significantly high relationship.

One strategy to assess the relative usefulness of two prognostic instruments is to compare their respective correlations with the criterion. On this basis marital status was as efficient as process-reactive scores. A second strategy is to determine the number of hits and misses in predicting length of future hospitalization. In employing this strategy one must select a cutting score for the ordinal measure (self-report scale) and compare the consequent hits and misses with the nominal categories of marital status. The conclusions from such a comparison must be limited since one should cross-validate cutting scores on a different population. A likely cutting point on the self-report scale which has been used for the assignment of Ss (Johnson & Reis, 1967; Meichenbaum, 1966b) is a median split, namely, those Ss who score 13 and above versus those who score 12 and below. Table 2 presents a comparison of the frequency distribution of self report scores and marital status for TDH 1-730 which has been divided into quartiles. An almost identical picture of efficiency is demonstrated for both process-reactive scores and marital status in predicting length of future hospitalization. Both measures are most efficient in predicting

who leaves the hospital earliest and who remains longest. Thus, using both the correlation with the criterion and number of misplacements, marital status is as efficient as the 24-item true-false self-report scale in predicting length of hospitalization.

Construct validity. Table 3 presents the intercorrelations between process-reactive scores, marital status, and cognitive measures. The significant negative correlations between premorbid adjustment, as assessed by the self-report scale, and marital status, and the similarities and proverbs test provide evidence for construct validity of these prognostic indicators. The present results are consistent with the findings of Ullmann and Eck (1965) who found a significant relationship between the self-report scale and cognitive performance as assessed by the Holtzman Inkblot Test. The finding that patients with poorer premorbid adjustment do poorer on cognitive tests was also found by Becker (1965), Little (1966), and Meichenbaum (1966a). The schizophrenic patients' scores on the proverbs, vocabulary, and similarities tests were not significantly related to any of the TDH measures ($N = 62$).

An analysis of cognitive performance by marital status revealed a consistent picture of married and divorced individuals both differing significantly from singles, but not from each other. Married and divorced patients performed at a superior level to single patients on the proverbs test and similarities subtest and on the Kent-Rosanoff word association

TABLE 2
FREQUENCY DISTRIBUTION ON TOTAL DAYS HOSPITALIZED AS A FUNCTION OF SELF-REPORT SCALE AND MARITAL STATUS

TDH	N	Self-report scale		Marital status	
		12 and below	13 and above	Married and divorced	Single
2-178	15	11	4	11	4
198-470	16	6	10	7	9
525-720	16	8	8	6	10
725-730	15	4	11	3	12

TABLE 3

INTERCORRELATIONS OF PROCESS-REACTIVE SCORES,
MARITAL STATUS, AND COGNITIVE MEASURES

Correlates	Marital status	Binet vocabulary	Similarities sub-test	Proverbs test
Process-reactive	.683**	-.100	-.225**	-.345**
Marital status		-.089	-.187*	-.323**
Vocabulary			.516**	.543**
Similarities				.708**

Note.— $N = 110$.* $p < .05$, one-tailed.** $p < .01$, one-tailed.

test,³ but they were not significantly different from each other. The married, divorced, and single patients did not significantly differ on the vocabulary test or on level of education, suggesting that level of intelligence or education does not explain differences in cognitive performance between groups.

Meichenbaum (1966b) tested 20 hospitalized male medical patients who had no history of psychiatric illness on these same cognitive measures. Of the 20 patients, 10 were married, 1 divorced, and 9 single. The medical patients who were single performed at a significantly inferior level ($p < .05$) on the proverbs test and gave less common associations on the Kent-Rosanoff word association test than married and divorced patients. As with the schizophrenics, there were no differences on vocabulary scores or level of education between single and nonsingle individuals.

In summary, a consistent negative relationship appears between level of premorbid adjustment, as assessed by self-report scale and marital status, and cognitive ability for both male schizophrenics and male normals. It does not appear that level of intelligence or education accounts for this relationship.

DISCUSSION

Evidence for the predictive validity of the self-report scale and marital status was found.

³ The Kent-Rosanoff word association test was administered only to the additional 48 schizophrenics (Meichenbaum, 1966b). A common association response was defined as one given by at least 100 of Russell and Jenkins (1954) Ss (see Ullmann, Krasner, & Edinger, 1964).

The correlation of these prognostic indicators and the criterion of length of hospitalization and the number of misplacements indicated that marital status was as efficient as the 24-item true-false process-reactive self-report scale. Garfield and Sundland (1966) have also found that marital status alone predicts length of hospitalization as well as prognostic scales designed to measure the process-reactive continuum. It appears that marital status (married, divorced, single) is as efficient in predicting prognosis as the prognostic scales (e.g., Elgin, Phillips, or self-report scales) thus far developed. Ullmann (1967) and Ullmann and Gurel (1964) found greater discriminability by dividing married patients into two groups, namely, those who list wives as people to notify in case of emergency and those who did not.

The reasons why marital status proves so effective in predicting prognosis are not known. Garfield and Sundland (1966) indicate two general sources of hypotheses which have been offered, namely, those which emphasize premorbid personality and those which stress social and environmental factors. In other words, there is something about the married individual which contributes to his leaving the hospital (e.g., he may be less ill to begin with), or the environmental pressures which surround the married individual contribute to his discharge. The data on the differences between married and single individuals on cognitive measures for both schizophrenic and nonschizophrenic males underscore the possibility that the married male may have a cognitive style which may contribute to better prognosis. A consistent and significant finding was that single males performed on an inferior level on proverbs and gave less common associations than married or divorced males, with intelligence held constant. A common characteristic of the performance of the single individuals on the proverbs test was an obsessiveness in giving responses. Comparing random samples of married and single males, singles gave significantly more words in their answers to proverbs resulting in more literal interpretations and an inferior performance. These results suggest that unmarried males in their late thirties, either schizophrenic or normal, are

more likely to manifest thought processes which may be described as ruminative, obsessive, and less common than marrieds. One does not know whether these thought processes contribute to the fact that males are not married or result from the absence of marriage. The heterosexual and familial responsibilities of marriage in most instances requires that the male use language in an instrumental manner, assert interpersonally, convey and validate thoughts and strategies, and make decisions. The unmarried male has less of an opportunity for interpersonal relationships and reciprocity of thoughts and thus is more prone to internalize decisions in the form of self-dialogue with the consequence of becoming more ruminative and obsessive. Thus, marital status may be seen as a reflection of a style of thought processes of the individual which correlates with amount of improvement and length of hospitalization. Further research needs to be conducted on cognitive styles of married versus single individuals in normal and maladaptive populations. The relationship between cognitive style and marital status in females may be quite different than in males.

The other major source of hypotheses designed to explain the efficacy of marital status as a prognostic indicator focuses on the social role and environmental situation of married and single individuals. These hypotheses suggest that married individuals leave the hospital because there are greater pressures for hospital discharge in order to return him to the community. At present we do not know if such pressures exist or how they influence the decision-making process. The emphasis in the development of prognostic scales has been on the instrument designed to measure the patient's premorbid adjustment and little attention has been paid to the factors that affect the criterion. An investigation of the judges (psychiatrists, psychologists, and social workers) may reveal why marital status plays such an important role in predicting prognosis. Stern, Stein, and Bloom (1956) have described a procedure one may follow to investigate the role marital status may play in influencing judges (see their discussion on analytic model).

In summary, an examination of the cogni-

tive style of married versus single males, and an investigation of the criterion of the judges may contribute to an understanding of why marital status is as efficient in predicting prognosis as more complex scales.

REFERENCES

- BECKER, W. C. A genetic approach to the interpretation and evaluation of the process-reactive distinction in schizophrenia. *Journal of Abnormal and Social Psychology*, 1956, 53, 229-237.
- BECKER, W. C. The process-reactive distinction: A key to the problem of schizophrenia? *Journal of Nervous and Mental Disease*, 1959, 129, 442-449.
- FARINA, A., GARMEZY, N., & BARRY, H. Relationship of marital status to incidence and prognosis of schizophrenia. *Journal of Abnormal and Social Psychology*, 1963, 67, 624-630.
- GARFIELD, S. L., & SUNDLAND, D. M. Prognostic scales in schizophrenia. *Journal of Consulting Psychology*, 1966, 30, 18-24.
- GARMEZY, N. Process and reactive schizophrenia: Some conceptions and issues. Unpublished manuscript, University of Minnesota, 1965.
- HELD, J. M., & CROMWELL, R. L. Premorbid adjustment in schizophrenia. *Journal of Nervous and Mental Disease*, 1968, 146, 264-272.
- HERRON, W. G. The process-reactive classification of schizophrenia. *Psychological Bulletin*, 1962, 59, 329-343.
- JOHNSON, M. H., & RIES, H. A. Validation study of the self-report scale for process-reactive schizophrenia. *Journal of Consulting Psychology*, 1967, 31, 321-322.
- KAUFMAN, L. N. The development of a proverb scale for the measurement of thinking pathology in schizophrenia and a further investigation of the process-reactive dimension. Unpublished manuscript, University of Illinois, 1960.
- LITTLE, L. N. Effects of interpersonal interaction on abstract thinking performance in schizophrenics. *Journal of Consulting Psychology*, 1966, 30, 158-164.
- MEICHENBAUM, D. H. Effects of social reinforcement on the level of abstraction in schizophrenics. *Journal of Abnormal Psychology*, 1966, 71, 354-362. (a)
- MEICHENBAUM, D. H. The effects of instructions and reinforcement on thinking and language behavior of schizophrenics. Unpublished doctoral dissertation, University of Illinois, 1966. (b)
- OBER, D. Reminiscence in schizophrenia. Unpublished manuscript, University of Illinois, 1966.
- PHILLIPS, L. Case history data and prognosis in schizophrenia. *Journal of Nervous and Mental Disease*, 1953, 117, 515-525.
- RUSSELL, W. A., & JENKINS, J. J. The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test. Tech-

- nical Report No. 11, 1954, University of Minnesota, Contract No. N8, ONR-66216, Office of Naval Research.
- STERN, G. G., STEIN, M. I., & BLOOM, B. S. *Methods in personality assessment*. Glencoe, Ill.: The Free Press, 1956.
- ULLMANN, L. P. *Institution and outcome*. Oxford: Pergamon Press, 1967.
- ULLMANN, L. P., & ECK, R. A. Inkblot perception and the process-reactive distinction. *Journal of Clinical Psychology*, 1965, 21, 311-313.
- ULLMANN, L. P., & GIOVANNINI, J. M. The development of a self-report measure of process-reactive continuum. *Journal of Nervous and Mental Disease*, 1964, 138, 38-43.
- ULLMANN, L. P., & GUREL, L. Size, staffing and psychiatric hospital effectiveness. *Archives of General Psychiatry*, 1964, 11, 360-367.
- ULLMANN, L. P., KRASNER, L., & EDINGER, R. L. Verbal conditioning of common associations in long-term schizophrenic patients. *Behavior Research and Therapy*, 1964, 2, 15-18.
- WITTMAN, P. Scale for measuring prognosis in schizophrenic patients. *Elgin State Hospital Papers*, 1941, 4, 20-33.
- ZIGLER, E., & PHILLIPS, L. Social competence and the process-reactive distinction in psychopathology. *Journal of Abnormal Psychology*, 1962, 65, 215-222.

(Received June 28, 1968)

(Continued from page 350)

- A-B Therapist "Types" and Responses to Patient-Communicate Hostility: An Analogue Study: Juris I. Berzins,* Edward Seidman, and Robert D. Welch: Department of Psychology, University of Kentucky, Lexington, Kentucky 40506.
- Individual Differences in Associating to Neutral and Emotional Words: W. P. Brown*: Department of Psychology, University of Aberdeen, Old Aberdeen, Great Britain.
- Examiner Sex, Subject Sex, and Rorschach Productivity: Sandra Harris* and Joseph Masling: Psychological Clinic, State University of New York at Buffalo, Buffalo, New York 14214.
- Effects of Stimulus Variation on Responses to the Group Version of the Holtzman Inkblot Technique: Wayne H. Holtzman,* Jon D. Swartz, and Jeffrey L. Sanders: Office of the Dean of the College of Education, University of Texas, Austin, Texas 78712.
- Effects of Sex Guilt and Sexual Stimulation on the Recall of Word Associations: Gary G. Galbraith* and Donald L. Mosher: Department of Psychology, Pennsylvania State University, 117 Burrowes Building, University Park, Pennsylvania 16802.
- Interpersonal Perception: Cognitive Complexity and Trait Implication: Charles F. Halverson, Jr.*: National Institute of Mental Health, Building 15-K, 9000 Rockville Pike, Bethesda, Maryland 20014.
- Use of Avoidance and Vigilance by Repressors and Sensitizers: Harvey E. Hoffman*: 1308 Willow Avenue, Apt. C-223, Melrose Park, Pennsylvania 19126.
- Strangers in Paradise: Robin C. Tucker*: Department of Psychology, Washington University, St. Louis, Missouri 63130.
- Test-Retest Characteristics of a Group of Attempted Suicide Patients: Alistair E. Philip* and J. W. McCulloch: Edinburgh University Department of Psychiatry, Royal Edinburgh Hospital, Morningside Park, Edinburgh 10, Scotland.
- Initiation of Cigarette Smoking: Is It Related to Parental Smoking Behavior?: Paul Wohlford*: Department of Psychology, University of Miami, Coral Gables, Florida 33124.
- Therapeutic Effectiveness of Untrained Volunteers with Chronic Patients: J. Scott Verinis*: Galesburg State Research, Galesburg, Illinois 61401.
- Dimensions of Outpatient Neurotic Pathology: Comparison of a Clinical versus an Empirical Assessment: Leonard R. Derogatis,* Ronald S. Lipman, Lino Covi, Karl Rickels, and E. H. Uhlenhuth: George Washington University, Biometric Laboratory, 1145 Nineteenth St., N.W., Washington, D. C. 10036.
- Sex-Role Stereotypes and Clinical Judgments of Mental Health: Inge K. Broverman,* Donald M. Broverman, Frank Clarkson, Paul Rosenkrantz, and Susan R. Vogel: Worcester State Hospital, Worcester, Massachusetts.
- Halo Effect in Examiner Scoring of Intelligence Test Responses: Jerome M. Sattler,* William A. Hillix, and Linda A. Neher: Department of Psychology, San Diego State College, San Diego, California 92115.
- Parental Perceptions of Young Children and Their Relation to Indexes of Adjustment: Emory L. Cowen,* Joan Huser, David R. Beach, and Julian Rappaport: Department of Psychology, University of Rochester, Rochester, New York 14627.

* Asterisk indicates author for whom address is supplied.

(Continued on page 364)

PATTERNS OF PERFORMANCE ON WISC SIMILARITIES IN EMOTIONALLY DISTURBED AND BRAIN-DAMAGED CHILDREN¹

LEON P. HALL² AND LAVERNE LADRIERE

Oakland Schools, Pontiac, Michigan

One hundred and twenty emotionally disturbed and 120 brain-damaged children in a public school setting were matched for WISC Full Scale IQ, CA, and sex. An analysis was made of the Similarities subtest responses using the error-analysis approach previously used for differential diagnosis of institutionalized adults' performance on the WAIS Similarities. The Wilcoxon matched-pairs signed-ranks technique was used in analyzing the data. Brain-damaged Ss made significantly more restrictive errors while emotionally disturbed Ss produced significantly more inadequate attempts at conceptualization. The discriminatory power of this error type achieved significance at the .01 level when comparing males alone, while it achieved only the .05 level of significance when applied to the female subgroups. Frequency of errors of the descriptive type and of positive correct responses did not distinguish between the two groups. Dynamic implications of findings were discussed.

Of the many attempts to differentiate various diagnostic subgroups on the basis of test performance, several studies have concentrated on the process of conceptualization, especially as related to perception of common elements or of relationships, as a possible critical diagnostic index. The work of Goldstein and Scheerer (1941) stressed the deficit in abstraction as observed in object sorting tasks as characteristic of brain-damaged individuals. Grassi (1953) distinguished between the performance of brain-damaged and of schizophrenic Ss on a block substitution test in terms of an abstract-concrete schema. McGaughran and Moran (1956) reported that conceptual communality or social agreement of concepts provided a criterion for discriminating between a schizophrenic and a nonpsychiatric population. A further study by McGaughran and Moran (1957) emphasized degree of abstraction as a major difference in the conceptual processes of brain-damaged patients who "underabstract" and schizophrenic patients who "overabstract." This last is closely related to Cameron's (1938) report of schizophrenic "overinclusiveness." The re-

search of Levanthal, McGaughran, and Moran (1959) analyzed the conceptual performance of brain-damaged, schizophrenic, and non-psychiatric groups on object sorting, block substitution, and similarities tests on the basis of simple-complex, public-private, open-closed, and abstract-concrete schematic axes.

Tutko and Spence (1962) noted that any approach involving analysis of correct responses primarily reiterates the already well-accepted premise that the hierarchy of response adequacy progresses from brain damaged to schizophrenic to normal, respectively, along an inferior-superior continuum. These authors further distinguished between two types of schizophrenic disturbance. The first of these is the process-type of schizophrenia in which onset is slow and insidious, premorbid social history is poor, and etiological factors may be somatogenic or organic in nature. The abstractive performance of these schizophrenics should approximate that of organic patients. In contrast is the reactive-type of schizophrenic syndrome in which onset is rapid and often related to some precipitating stressful event, premorbid social history is adequate, and etiology is assumed to be primarily psychogenic in nature. Abstractive performance of this group should follow that attributed to schizophrenic populations usually described in the literature. Tutko and

¹The authors are grateful to Anne Anastasi for her critical reading and helpful suggestions in the preparation of this study.

²Requests for reprints should be sent to Leon P. Hall, Oakland Schools, Campus Drive, County Service Center, Pontiac, Michigan 48053.

Spence found significant differences in performance on the Goldstein-Scheerer Sorting Test between brain-damaged Ss and process schizophrenics whose protocols contained a greater percentage of restrictive errors (i.e., descriptive, nonconceptual, stimulus bound, concrete responses) and reactive schizophrenics who presented protocols yielding the highest percentage of expansive errors (i.e., over-inclusive, overabstractive, poor, vague, personal responses) of all groups tested. Their study used a modified form of error analysis based on a previous study (Rapaport, Gill, & Schafer, 1945). In a more recent study, Spence (1963) extended this approach to a comparative analysis of incorrect responses on the WAIS Similarities subtest of schizophrenic, organic, and normal adult males. Significant differences were in the expected direction, with the organic group giving the greater number of restrictive errors. Additional data revealed that the protocols of the brain-damaged group also showed the highest frequency of "Don't Know" responses, whereas the schizophrenic group tended to avoid this response.

Korman's (1960) recent study has indicated that the type, extent, and location of damage is directly related to psychological performance in organic patients. Similarly, the importance of sample homogeneity in comparisons of behavioral dynamics between different diagnostic groups of more psychogenic origin is underlined by distinctions such as the one made by Spence (1963) herself between reactive and process schizophrenics.

Stimulated by Spence's findings, Watson (1965) analyzed error-type frequencies across 10 WAIS subtests (Digit Symbol excluded) in the protocols of 40 male schizophrenic and 40 male organic patients. Watson questioned the usefulness of this approach in differential diagnosis not only due to the fact that only 3 of the 40 possible differences between groups achieved significance, but also due to the progressive disappearance of any such differences as a direct function of length of hospitalization. Watson's recategorization of error types as well as his application of these error groupings to the aggregate of 10 WAIS subtests instead of to the Similarities subtest on which they were developed reduced

the comparability of his data with that of Spence (1963) and obscured the dynamic rationale of Spence's analysis.

The present study was initiated to examine further the diagnostic potential of Spence's approach. In order to assure comparability of data, error analysis in the present study was restricted to the Similarities subtest alone and all of Spence's original error groupings were used. Most previous studies using error analysis had restricted themselves to adult male Ss. It was felt that an evaluation of the generality and validity of Spence's approach could be best served by its application to a younger noninstitutionalized population. Therefore Ss used in the present study range in age from 6 to 15 years. The matched-pair technique was used to control the factors of age, sex, and intellectual ability. With these observations in mind, it was hypothesized that a significantly greater proportion of restrictive errors would occur in the protocols of the organic sample; whereas expansive errors were expected with significantly greater frequency in the responses of the non-organic, psychiatric sample. Finally, comparisons within syndromes were undertaken to determine the extent of influence of such factors as chronological age, level of measured intelligence, and sex on the diagnostic value of error analysis.

METHOD

Subjects

WISC protocols were obtained on Ss referred to the Psychological Services of Oakland Schools. The organic youngsters (BD) were diagnosed on the basis of evidence of probable physiological trauma due to injury and/or illness in life histories as well as developmental and behavioral signs of organicity. Many of these records also included neurological and/or EEG data or supportive medical diagnosis. The non-organic psychiatric Ss (ED) were diagnosed on the basis of parallel criteria including bizarre and/or abnormal behavior, disturbed development, and history of psychologically traumatic circumstances. Often these records were supplemented with psychiatric diagnoses of serious disturbance. In no case were psychometric signs of either organicity or emotional pathology used for sample selection. All youngsters used in this study were public school referrals and, although many were under medical and/or psychiatric care, none had been institutionalized up to the time of testing.

It should be noted that the group in the present study is referred to not as schizophrenic but as either

emotionally disturbed or as the nonorganic psychiatric sample. This is not to equate the diagnosis of adult schizophrenia with that of serious emotional disturbance in children. Further, since the syndrome classified as schizophrenia in adults has much broader implications and is applied to a considerably wider range of symptoms and dynamics than that of childhood schizophrenia, it was felt that the emotionally disturbed sample provided a better approximation of the adult reactive schizophrenic psychiatric sample in the literature. This assumption was enhanced by the fact that the great majority of youngsters included in the emotionally disturbed sample were characterized by symptoms of withdrawal, bizarre behavior, and social isolation. They were diagnosed as severely disturbed with no known organic symptoms or history.

Sampling Procedures

The Ss were drawn from a data pool of 342 referrals which had been definitely diagnosed as either organic or emotionally disturbed. Of these, 171 children (103 boys and 68 girls) were emotionally disturbed, while 171 children (113 boys and 58 girls) were organics. The matched-pair technique was used in selecting the following subgroups from this population for consideration:³

1. One hundred and twenty emotionally disturbed and 120 organic youngsters, matched for WISC Full Scale IQ, chronological age, and sex, were compared in terms of relative frequency of error type across syndrome. This constituted the sample from which the major conclusions derived from the hypothesis were drawn. To permit further comparisons with previous data on the male population alone, both the organic and the emotionally disturbed samples were broken down according to sex into comparative subgroups of 80 boys and 40 girls (the other matching data remaining constant).

2. Since the secondary purpose of the present study constituted an evaluation of the relative influence of chronological age, sex, and ability on the diagnostic value of the error analysis approach, the following subgroups were pulled from the data pool to accomplish this purpose:

2a. Sex Differences: 40 organic boys were compared with 40 organic girls, matched for chronological age and Full Scale IQ, in terms of Similarities error frequencies. Similarly, 40 emotionally disturbed boys were compared with 40 emotionally disturbed girls, again paired according to chronological age and Full Scale IQ, regarding various error frequencies.

2b. IQ Differences: The relative error frequencies in the protocols of 40 organic boys of average-to-low

normal intellectual ability (i.e., Full Scale IQ of 80 or better) were compared with those of 40 organic boys of poor-to-retarded intellectual ability (i.e., Full Scale IQ of 79 or less). Similar comparisons using the same cut off of Full Scale IQ of 80 were made of error frequencies in protocols of 40 average-to-slow versus 40 poor-to-retarded emotionally disturbed boys. In both comparisons, youngsters were paired according to sex and chronological age, but were deliberately chosen to be as far apart as possible in intellectual ability.

2c. Chronological Age Differences: WISC Similarities error frequencies were compared in the protocols of 40 organic boys of 11½ years or older versus those of 40 organic boys under that age. Similar comparisons were made between responses of 40 emotionally disturbed boys above 11½ years and 40 emotionally disturbed boys under 11½ years. All comparisons were based on youngsters paired in terms of WISC Full Scale IQ and WISC Similarities scores as well as on sex; however, youngsters were selected to make the age gap as wide as possible.

Scoring Procedures

Incorrect responses (no-credit) on the WISC Similarities Items 5 through 16 (Items 1-4 are analogies, not similarities, and were, therefore, excluded from scoring) were relegated into one of six basic error categories, following Spence's (1963) outline:

1. Inadequate Conceptual (C-) attempts specification of common properties that are too vague, inaccurate, or idiosyncratic to be considered correct.
2. Narrative and Descriptive (ND) responses concerning one or both of the items separately, but with no specification of common properties.
3. Don't Know (DK) responses in which S explicitly indicates he did not know how the two items were similar.
4. No Response (NR) given to the question.
5. Don't Know-No Response (DKNR).⁴ This error category includes both the explicit Don't Know responses noted in Category 3 and the Don't Know responses implicit in the No Response or Category 4 error type.
6. Denials (De) of similarity.

All protocols were scored by two psychologists independently. Of the 1,085 total items scored, disagreement occurred in only 19 (1.25% of the cases).

⁴ Tutko and Spence (1962) included both narrative and descriptive errors (the present study's Category 2) and no-response categories (4 and 5) in their definition of restrictive-type errors. The same definition of restrictive errors was retained herein. Thus, restrictive errors were operationally defined in terms of Error Categories 2 and 5 (the latter including the DK errors as part of the NR group), again attempting to parallel the Tutko and Spence (1962) work.

³ The complete matching data for all comparisons made in this article are available in tabular form on request to the authors. Tests for biasing influences of sample heterogeneity and regression effects due to inclusion of IQ as one of the matching criteria have been evaluated. This information is also available on request.

RESULTS AND DISCUSSION

The major purpose of the present study was to extend Spence's error analysis approach by evaluating its diagnostic potential when applied to children of both sexes. It was hypothesized that expansive errors (the present study's Category 1, represented in Spence's analysis as the C— error category) would occur with significantly greater frequency in the protocols of children diagnosed as emotionally disturbed while restrictive errors (related by Spence to the ND and DK-NR error categories) would be made significantly more often by children diagnosed as organic. Table 1 presents the frequency of occurrence of each error category of matched pairs for both diagnostic groups as well as the comparisons which significantly discriminated between these diagnostic groups as determined by application of the Wilcoxon matched-pairs technique. Because most of the literature has presented data solely on male populations, it was felt that it would be meaningful to present a comparison of the frequency of occurrence of each error type between diagnostic groups for boys and for girls within those groups separately. This would not only allow greater comparability with the literature by providing an exclusively male population, but also presents the opportunity to evaluate the differentiating potential of error analysis of this type in terms of female samples. These data, too, are included in Table 1.

Perhaps the most outstanding finding re-

vealed in Table 1 is that the C— error occurs significantly more often in the emotionally disturbed group than in the organic group, both when sexes are combined and for each sex considered separately. This directly confirms the hypothesis of the present study and offers strong support for Spence's findings concerning the diagnostic value of this type of error. It also conforms to the generally accepted relationship of the expansive error to nonorganic psychiatric groups. Perhaps one generalization may be that the vague, inappropriate, idiosyncratic, and conceptually inadequate responses of the C— or expansive type reflect loss of reality contact, intrusion of personal (often autistic) material into the thought process, and the overresponsive, over-abstractive behavior characteristic of severe psychogenic disturbance.

The only other error category in which error frequency distinguished significantly between the two syndromes is the restrictive error, DK. This error occurs with significantly more frequency in the protocols of organic boys than in those of their psychiatric counterparts. The tendency represented by this type of error is consistent with the literature and is generally interpreted to reveal the diagnostic characteristics of the organic, notably, impairment of abstractive ability, concreteness, and deficient conceptualization. However, any interpretation of this kind must be restricted by the fact that the DK error appears with close-to-equal frequency in the protocols of emotionally disturbed and of organic girls alike. It fails to discriminate between them to such an extent that when the data of the sexes for each diagnostic group are combined only a moderate tendency for organics to make more DK errors is observed. The inclusion of the NR error frequencies with those of the DK category, as done by Spence, appears only to obscure the limited diagnostic value of the DK response. It would seem safe to assume that the NR error dynamically represents something other than an implicit Don't Know and should not be arbitrarily combined with it. The rationale for the specific separation of this combined classification into its two component error types (apparently not done by Spence or Watson) lay in the direct assumption in the

TABLE 1

FREQUENCY OF OCCURRENCE OF THE DIFFERENT
ERROR TYPES BETWEEN DIAGNOSTIC GROUPS

Group	N	Error category					
		C—	ND	DK	NR	DKNR	De
Total BD	120	142 ^{**}	110	136	30	166	23
Total ED	120	180 ^{**}	76	116	35	151	24
BD boys	80	111 ^{**}	72	94 [*]	19	113	11
ED boys	80	134 ^{**}	47	71 [*]	21	92	20
BD girls	40	31 [*]	38	42	11	53	12
ED girls	40	46 [*]	29	45	14	59	4

* $p < .05$.

** $p < .01$.

literature that the specific DK response was dynamically related to the organic syndrome while the NR response might have represented something other than an implicit DK (e.g., negativism) and, in so doing, might have contaminated the DK category, diluting possible significant differences between groups inherent in the DK classification.

It is also obvious from Table 1 that the ND error category fails to discriminate significantly between the organic and the psychiatric Ss in the present study. This agrees with Spence's (1963) data which also revealed no significant differences between her two diagnostic groups in the frequency of the ND response. Since the ND error has usually fallen into the restrictive category as a chaining or naming response, stimulus bound and concrete, Spence had initially anticipated its greater frequency of occurrence in her organic group. Its failure to discriminate between the two groups suggested the necessity of a broader dynamic interpretation of this type of incorrect response. It may indeed have represented the stimulus bound, nonintegrative, perseverative naming and chaining process characteristic of the organic mental process. However, it may have been equally indicative of the overinclusiveness process often noted in the nonorganic psychiatric patient. Theoretically, at least, this second possibility would allow for equal occurrence of this kind of error in both syndromes. However, a closer examination of the frequency data of the present study does not indicate that this is quite the case. It is easily seen that there is a strong tendency for greater frequency of the ND response among organics both when the groups are combined or when the sexes are considered separately. It would seem that this error category has potential diagnostic significance. Perhaps further refinement in scoring procedures to separate the perseverative naming of the organic from the combination responses characteristic of psychiatric patients might eliminate the overlap and realize this potential.

The final error type for consideration is the relatively rare De response. Spence relates this type of error to several possible dynamics. First, it could be related to the schizophrenic's indifference to reality stan-

dards, his unwillingness to give a DK response, and, in such a frame of reference, might be a downward extension of the C-response. Second, it might represent the suspiciousness in the schizophrenic personality such as was noted by Weiner (1957). Finally, it might be simply negativism. In all of these interpretations the De response dynamically fits the psychiatric patient, and Spence's data showed that although rare and therefore statistically tentative due to small *N*, for her sample, the De response occurred significantly more frequently in the protocols of schizophrenics than those of organics. The present data fail to confirm this finding, nor do they reveal any tendency in that direction. Further, on the assumption that the NR response might as easily represent negativism or psychotic muteness as the implicit Don't Know assumed by Spence, the De and NR categories were combined, to the advantage of neither. It might be asserted that due to the relative rarity and to the general failure to discriminate diagnostically of the NR error and possibly of the De category, both are of little value in the general application of error analysis.

In order to better define the relationships indicated in Table 1 the conduct of this research turned to an analysis of the influence of the variables of chronological age, sex, and intellectual ability on the diagnostic value of the error-analysis approach. Comparisons were made within each diagnostic syndrome to test the effect of these parameters on the frequency with which the various types of errors occurred. Again, the matched-pairs technique was applied in the selection of these subgroup samples from the data pool. The Wilcoxon test was used to measure the significance of the differences in error frequencies between comparison groups. The frequency data for these subsidiary comparisons can be found in Table 2 along with indications of significance wherever it appeared.

The first set of comparisons attempted to determine whether or not sex was a major factor influencing error frequencies. As can be seen from Table 2 the comparison of error responses of organic boys versus organic girls yielded no significant differences in any error type between the sexes. Similar comparison

TABLE 2

COMPARISON OF THE INFLUENCE OF SEX, IQ, AND
CHRONOLOGICAL AGE ON ERROR FREQUENCIES
WITHIN DIAGNOSTIC GROUPS

Group	N	Error category					
		C—	ND	NR	DK	DKNR	De
BD male	40	45	31	13	42	55	7
BD female	40	33	35	15	41	56	11
ED male	40	78	32	12	30	42	8
ED female	40	44	22	26	46	72	1
BD male—low IQ	40	32	42	9	37	46	4
BD male—high IQ	40	56	26	6	61	67	0
ED male—low IQ	40	71	34	6	40	46	9
ED male—high IQ	40	67	24	12	56	68	3
BD male—older	40	67	32	10	42	52	5
BD male—younger	40	33	31	9	45	54	2
ED male—older	40	68	26	12	44	56	9
ED male—younger	40	68	32	14	33	47	6

* $p < .05$.

** $p < .01$.

of error frequencies in protocols of emotionally disturbed boys versus those of emotionally disturbed girls revealed significant differences between the sexes in the frequency of both C— and DKNR error types. Examination of the frequency data showed that the emotionally disturbed boys made nearly twice as many C— responses as did their feminine counterparts, while the emotionally disturbed girls gave significantly more restrictive errors of both the DK and the NR type. These findings reflected those presented in Table 1 in that although the C— response occurred with significantly greater frequency in the protocols of the emotionally disturbed Ss than in those of the organic sample, regardless of sex, the discriminatory power of this error type achieved only the .05 level of significance when applied to the female subgroups. These data also help to explain why the restrictive error category of DK and DKNR failed to discriminate between psychiatric and organic groups with the one exception of the DK error type which occurred significantly more frequently among organic than among emotionally disturbed boys. It would seem that while emotionally disturbed girls did reflect their disturbance in their significantly greater use of C— responses than did their organic pairs, they also tended to resort with significantly greater frequency to the DK and DKNR errors than did boys within the same syndrome. The restrictive errors of the psychiatric girls equalled the frequency with which

the organic females responded thus masking the full meaning of the diagnostic potential of the DK error. This was not true when the sexes were considered separately. These results suggest greater homogeneity as well as greater response determination by syndrome dynamics among the emotionally disturbed boys than among girls within the same syndrome. Further they again call into question, at least among female samples, the assumption that DK and NR errors must be restrictive errors representing organic concreteness, and present the alternative possibility, mentioned above, that such responses might also reflect the negativism and/or autistic muteness of the psychiatrically disturbed youngster.

Comparisons were then made within each diagnostic syndrome separately to investigate the influence of chronological age and intellectual ability on the frequency of occurrence of various error types. Due to limited number of girls available, these comparisons were based on all male samples. These data also appear in Table 2.

When responses of brain-damaged boys of average-to-dull normal intelligence compared with those of brain-damaged boys of equal age but low intelligence, it was noted that the C— error occurred with greater frequency in the protocols of the more intellectually able youngsters. Similarly, older organic boys gave significantly more C— responses than did their younger counterparts when matched for ability. This seemed to indicate that within the organic group, the C— error is not only a function of intelligence but also of age. Thus, brighter and/or older and more experienced youngsters tended to attempt both the generalization and greater adequacy of response that the C— error category implies regardless of their organic limitations than do younger or more retarded organic Ss.

This position that syndrome dynamics might be considerably influenced by age and ability factors is further supported by the fact that the brighter organic boys gave significantly more DK responses than did the more retarded group, while the low-ability group tended to produce significantly more ND responses. If the DK and ND responses do, as assumed throughout the literature, represent dynamically the stimulus-bound and

perseverative behavior characteristic of organicity alone, then no significant differences within the organic syndrome should be observed. Thus, the fact that the brighter organic group gave significantly more DK responses than the less able group must suggest the importance of ability as well as syndrome dynamics in the frequency with which this error occurred. This finding could be interpreted along the same line as was the greater frequency of the C— error with brighter youngsters. Thus, it would seem that the brighter organic youngster not only attempted to make more adequate and abstract responses but also was more aware of what he did and did not know than his slower counterparts.

It might be argued that these ability-related differences merely represent the general tendency of bright youngsters to give more responses of all kinds than slower youngsters. This argument would have appeared a great deal stronger had not the retarded organic group given almost twice as many ND responses than the more able group. The brighter youngster tries for responses with greater generality and possesses greater awareness of reality and personal limitation; whereas the duller organic youngster evidenced the fragmented and perseverative and stimulus-bound responses assumed to be universally characteristic of organicity. Thus, it would seem that the diagnostic value of error analysis relative to organics is not simply a direct function of syndrome dynamics but is also related to age and ability where the C— and DK error types are concerned, and represents an inverse function of ability in regard to the ND error types.

A strong contrast to the above was apparent when similar comparisons of the effect of ability and age variables on error frequencies in the protocols of emotionally disturbed boys were made. These data, also presented in Table 2, reveal no significant differences relative to chronological ages and only one significant difference relative to ability in error frequencies across the error categories under consideration. This suggests not only greater dynamic homogeneity within the emotionally disturbed sample, but also greater overall influence of syndrome dynamics on error frequency patterns than was observed in the

protocols of the organic group. This interpretation of response determination based primarily on syndrome dynamics is underlined by the uniformly high occurrence of the C— error type across age and ability subgroups within the emotionally disturbed sample since the C— error type is the one most associated with psychiatric disturbance. The single case in which significant differences relative to ability occurred was that of the error category DKNR. Responses of this type were given more often by brighter emotionally disturbed boys. As in the above, this cannot be related simply to greater overall response frequency in protocols of brighter children since the frequency data do not confirm this. Rather, it might be suggested that the brighter sample of emotionally disturbed boys use this type of response as an indication of either better awareness of personal limitation and reality or of greater negativism, suspiciousness, and resistance, than is existent in the less intellectually able emotionally disturbed boys.

SUMMARY AND CONCLUSIONS

In summary, the data of the present study generally support the validity and diagnostic usefulness of the error analysis approach. The fact that the expansive (or C—) error type occurred with significantly greater frequency in the responses of the psychiatrically disturbed group than in those of the organic group regardless of sex provides strong confirmation of Spence's (1963) findings and of the hypothesis of the present research based thereon. The diagnostic merit of this type of error appears even more powerful when it is noted that it is not test-specific, but equally effective whether measured on the WAIS or the WISC. It is concluded that this error category is primarily a reflection of syndrome dynamics with diagnostic potential. Strong trends in the predicted direction of occurrence appeared in the data with regard to the restrictive-type error category (i.e., DK, NR, DKNR), and the DK error frequency did discriminate significantly between organic versus emotionally disturbed boys. It is suggested that these findings support those of Spence, but the diagnostic value of this type of error is of a more limited nature than that of the C— category. The restrictive error

type ND revealed a mixed picture suggesting that this error category needs further consideration and refinement before its diagnostic value can be validly assessed. Finally, differential effects of chronological age, IQ, and sex were analyzed and discussed not only in terms of their direct influence on error analysis but also relative to their added interpretative value and restrictions in the use of this approach to differential diagnosis.

REFERENCES

- CAMERON, N. Reasoning, regression and communication in schizophrenics. *Psychological Monographs*, 1938, 50 (1, Whole No. 221).
- GOLDSTEIN, K., & SCHEERER, M. Abstract and concrete behavior: An experimental study with special tests. *Psychological Monographs*, 1941, 53(2, Whole No. 239).
- GRASSI, J. R. *The Grassi Block Substitution Test for measuring organic brain pathology*. Springfield, Ill.: Charles C Thomas, 1953.
- KORMAN, M. Some problems in the psychological diagnosis of brain damage: An overview. *Texas Reports on Biology and Medicine*, 1960, 18, 214-221.
- LEVENTHAL, D. B., MCGAUGHRAN, L. S., & MORAN, L. J. Multivariable analysis of the conceptual behavior of schizophrenic and brain-damaged patients. *Journal of Abnormal and Social Psychology*, 1959, 58, 84-90.
- MCGAUGHRAN, L. S., & MORAN, L. J. "Conceptual level" vs. "conceptual area" analysis of object sorting behavior of schizophrenic and nonpsychiatric group. *Journal of Abnormal and Social Psychology*, 1956, 52, 43-58.
- MCGAUGHRAN, L. S., & MORAN, L. J. Differences between schizophrenic and brain-damaged groups in conceptual aspects of object-sorting. *Journal of Abnormal and Social Psychology*, 1957, 54, 44-49.
- RAPAPORT, D., GILL, M., & SCHAFER, R. *Diagnostic psychological testing: The theory, statistical evaluation, and diagnostic application of a battery of tests*. Vol. 1. Chicago: Yearbook Publishers, 1945.
- SPENCE, J. T. Patterns of performance on WAIS Similarities in schizophrenic, brain-damaged and normal subjects. *Psychological Reports*, 1963, 13, 431-436.
- TUTKO, T. A., & SPENCE, J. T. The performance of process and reactive schizophrenics and brain-injured subjects on a conceptual task. *Journal of Abnormal and Social Psychology*, 1962, 65, 387-394.
- WATSON, C. G. WAIS error types in schizophrenics and organics. *Psychological Reports*, 1965, 16, 527-530.
- WEINER, G. The effect of distrust on some aspects of intelligence test behavior. *Journal of Consulting Psychology*, 1957, 21, 127-130.

(Received July 1, 1968)

(Continued from page 356)

- Predictive Validity of Six Neurodiagnostic Tests: A Decision Theory Analysis: Paul Satz,* Eileen Fennell, and Carol Reilly: Department of Clinical Psychology, University of Florida Medical Center, Gainesville, Florida.
- Motivational Properties of Personality Conflict: Ronald E. Smith*: Department of Psychology, University of Washington, Seattle, Washington 98105.
- Perceived Maternal Nurturance and Control of Process Schizophrenics, Reactive Schizophrenics, and Normals: Peter A. Magaro* and Barbara A. Hanson: Department of Psychology, University of Maine, Orono, Maine 04473.
- Performance Characteristics of School Children on the BIP Bender Test: Jerry Adams and Arthur Canter*: Department of Psychiatry, State Psychopathic Hospital, 500 Newton Road, Iowa City, Iowa 52240.

* Asterisk indicates author for whom address is supplied.

AN ATTEMPT TO TEST ASSUMPTIONS ABOUT SOME INDICATIONS OF NEGATIVISM ON PSYCHOLOGICAL TESTS

ELIZABETH FOX AND SIDNEY J. BLATT¹

Yale University

Two cognitive-perceptual measures on psychological tests, one from the WAIS and the other from the Rorschach, are frequently assumed to reflect tendencies toward oppositionality and negativism. Oppositional behavior is considered when the number of digits remembered in a backward order exceed the number of digits remembered in a forward order or when there are White Space responses on the Rorschach. A significant relationship was found between these two independent measures and this finding was discussed in relation to the validation of interpretative assumptions of psychological tests.

On the Wechsler Digit Span, the usual pattern is for a person to remember more digits in forward order than they remember in reverse order. Rapaport, Gill, and Schafer (1945, p. 180) suggest that a pattern of a greater number of digits remembered in a reverse order rather than in a forward order may reflect, among other things, negativism and a tendency to be oppositional. In the Rorschach literature, there is the assumption that a tendency toward negativism and oppositionality can also be reflected in the White Space (*S*) responses. Rorschach (1942) believed that *S* responses always indicated "oppositional tendency whether they occurred in large or small numbers. In normals they indicate argumentative, willful, obstinate and querulous types; in schizophrenics they indicate blocking, negativism and eccentricity [p. 40]." Schafer (1954) in discussing obsessive-compulsive defensive operations comments that "the underlying negativistic, rebellious, stubborn impulses of persons characterized by rigid reaction formations against hostility frequently find expression through *S*, . . . *R* and *Dr* may increase greatly . . . either high or low *R* may result from hostile impulses and demanding attitudes . . . [p. 353]." Rapaport, Gill, and Schafer (1946, p. 178) point out that space and *Dr* responses tend to accumulate in the same records. They also indicate that the difficulty in validating the assumptions about test responses, which indi-

cate oppositional tendencies, is a function of the many alternative ways such tendencies can occur in behavior and in test records (1946, pp. 176-177).

Several research strategies could be used to test these assumptions about indicators of negativism in the various psychological tests. Ideally, test variables might be studied in relation to overt behavior, but it is often difficult to obtain reliable measurements of the degree of negativism in overt behavior that occurs in a complex social matrix or in an experimental procedure. At least as a first step, however, one can test assumptions about the psychological test indications of negativism by examining the relationship between the balance of digits forward (*DF*) to digits backwards (*DB*) and Rorschach *S* responses.

METHOD

From a file of approximately 700 psychological test records, all patients who had *DB* in excess of *DF* were selected for study. The diagnoses of these patients ranged from psychoses through neuroses but records were excluded if there were suggestions of central nervous system impairment. The experimental group consisted of 14 records (7 males and 7 females) and two control groups were matched with the experimental group for age, sex, and Full Scale IQ. The first control group consisted of *Ss* with *DB* = *DF* and the second control group consisted of *Ss* where *DB* was less than *DF* by one or two digits. These three groups were compared on a number of Rorschach variables and WAIS subtest scores.

RESULTS

As presented in Table 1, the data indicate that the number of *S* and the number of *Dr*

¹Requests for reprints should be sent to Sidney J. Blatt, Department of Psychiatry, Yale University, 333 Cedar Street, New Haven, Connecticut 06510.

TABLE 1

RELATIONSHIPS OF THE BALANCE OF DIGITS
FORWARD AND BACKWARD TO WAIS
SUBTEST SCORES AND RORSCHACH
VARIABLES

Variables	DB > DF	DB = DF	DB < DF	F ^a
Control				
Age	27.0	24.8	26.4	.95
Full Scale IQ	109.00	108.00	110.21	.95
WAIS				
Comp	12.50	11.43	10.50	1.88
Info	11.29	11.21	11.29	.00
D Sp	11.86	11.00	9.93	1.25
Arith	10.79	10.57	9.86	.27
Simil	10.86	10.64	10.50	.07
Vocab	11.86	11.93	11.71	.02
PA	11.21	10.57	10.86	.20
PC	10.57	9.86	11.64	2.39
BD	11.21	11.00	11.71	.31
OA	10.86	10.43	11.86	.99
D Sy	9.71	10.79	11.29	1.53
Verbal IQ	110.00	108.57	107.29	.14
P IQ	106.36	105.71	112.14	1.53
Rorschach				
R	34.21	27.71	25.57	1.78
W	6.64	8.14	9.71	1.17
D	18.93	16.57	12.50	1.69
Dd	.71	.14	.57	.73
Dr	5.50	2.07	1.57	8.33**
Do	.21	.36	.14	.42
De	.36	0	.14	1.04
S	1.29	.29	.50	3.67*
S (+ additionals)	3.36	1.85	1.00	3.28*
s	.36	0	.21	1.00
DW	.07	.14	.14	.15
M	1.57	1.36	2.00	.76
2C	3.86	3.14	3.89	.22
F ₀ %	66.50	70.64	61.64	1.08
F ₀ % Exl	91.07	93.07	86.00	1.77
F + %	67.14	70.50	69.57	.19
F + % Exl	67.71	71.86	67.64	.47

Note.—N = 14 for each group.

^a df = 2/39.

* $p < .05$ ($F = 3.24$).

** $p < .01$ ($F = 5.21$).

responses relates to the balance of DF to DB. No differences were found between the three groups on any WAIS subtests or on traditional Rorschach variables other than the number of *S* and *Dr* responses.

DISCUSSION

The significant relationship between two independent cognitive-perceptual measures which were assumed to reflect oppositionality and negativism tends to support the hypotheses about the psychological processes involved in *S* responses and the balance of digits backward to digits forward. It is also consistent that the groups should also differ significantly on *Dr* responses since Rapaport et al. (1946) commented that *Dr* and *S* responses tend to co-occur.

Though there is a significant relationship between these independent cognitive-perceptual measures which are assumed to reflect negativism and oppositionality, support

for the assumption about test indicators of negativism can only be inferential since no direct assessment was made of manifest overt behavior. Yet the expectation that there would be a relationship between the variables on the Wechsler and the Rorschach was based on independent clinical observation and theory. The relationship of test variables to particular psychological processes may require a demonstration of a relationship to manifest behavior but the relationship between internal psychological states and overt behavior is a complex and difficult problem. There is little evidence of a one-to-one, isomorphic relationship between internal psychological states and specific overt behavior (Skolnick, 1966). Psychological processes are only one determinant of behavior in a complex social matrix (Engel & Blatt, 1963) and though we may understand indications of psychological states and tendencies toward certain types of behavior, we may not be able to predict how these internal states will specifically manifest themselves in social situations. The more we attempt to place constraints and controls on the variations of the social matrix so there is a limited range of possible behavior, the more the behavior moves toward becoming only responses to other psychological test procedures. The correlation of dimensions on independent psychological procedures, as in this study between the WAIS and Rorschach, may be one way to test assumptions about the interpretation of psychological test variables.

REFERENCES

- ENGEL, M., & BLATT, S. J. Clinical inference in psychological assessment. Paper presented at the meeting of the American Psychological Association, Philadelphia, September 1963.
- RAPAPORT, D., GILL, M., & SCHAFER, R. *Diagnostic psychological testing*. Vol. 1. Chicago: Year Book Publishers, 1945.
- RAPAPORT, D., GILL, M., & SCHAFER, R. *Diagnostic psychological testing*. Vol. 2. Chicago: Year Book Publishers, 1946.
- RORSCHACH, H. *Psychodiagnostics*. (Trans. by P. Lemkau & B. Kronenberg) New York: Grune & Stratton, 1942.
- SCHAFER, R. *Psychoanalytic interpretation in Rorschach testing*. New York: Grune & Stratton, 1954.
- SKOLNICK, A. Motivational imagery and behavior over twenty years. *Journal of Consulting Psychology*, 1966, 30, 463-478.

(Received July 15, 1968)

SENSATION SEEKING SCALE AS A PREDICTOR OF NEED FOR STIMULATION DURING SENSORY RESTRICTION

JOHN HOCKING¹ AND MALCOLM ROBERTSON

Western Michigan University

The study tested the hypothesis that the Zuckerman Sensation Seeking Scale (SSS) predicts strength of need for stimulation under conditions of sensory restriction. During a 3-hour isolation period, 15 high- and 15 low-SSS scorers were free to button press for visual, auditory, and kinesthetic stimulation. Except for visual stimulation high-SSS scorers button pressed at a higher rate than low-SSS scorers, but the differences were not significant. Low-SSS scorers showed a significantly greater need for visual stimulation than high-SSS scorers. High-SSS scorers showed a significant preference for kinesthetic stimulation, while low-SSS scorers had a significant preference for visual stimulation. Results were discussed in the light of future research strategy.

A number of writers have discussed the concept of optimal stimulation, by which they mean that under conditions of sensory deprivation or overstimulation the organism seeks to increase or decrease the sensory input to a level conducive to optimal functioning (Duffy, 1962; Fiske & Maddi, 1961; Hebb, 1955; Leuba, 1955; Schultz, 1965). With few exceptions, experimental work has been directed toward the organism's behavior in response to conditions of restricted sensory input (Ansfield, 1965; Bexton, 1954; Goldstein, 1965; Vernon & McGill, 1960; Zuckerman & Haber, 1965). One major research limitation has been the lack of base-line data on an individual's level of optimal stimulation. To correct for this deficiency, Zuckerman, Kolin, Price, and Zoob (1964) developed the Sensation Seeking Scale (SSS) to provide a quantitative measure of the strength of need for stimulation. The SSS is a forced-choice questionnaire in which the items, derived from factor analysis, pertain to preferences for intensities of sensation (heat, cold, noise, taste, colors), familiarity as opposed to novelty, routine as opposed to irregularity, and security as opposed to adventure. Zuckerman hypothesized that an individual who scores high on the SSS would have a strong need for stimulation and would seek a relatively high level of sensory input.

Conversely, an individual who scores low on the SSS would have less need for stimulation and would seek a relatively lower level of sensory input.

The present study tested the hypothesis that the SSS predicts strength of need for stimulation in a sensory restricted environment. Specifically, high-SSS scorers were expected to request more stimulation than low-SSS scorers during 3 hours of sensory restriction.

METHOD

Sample and Apparatus

The SSS was administered to 214 male undergraduates, and the 15 highest and lowest scorers comprised the sample. Out of a maximum score of 26 on the SSS, the mean of the high scorers was 23 ($SD = .90$), and the mean of the low scorers was 8 ($SD = 1.8$). Approximately 1-3 weeks later, each *S* was retested with the SSS ($\rho = .93$), and then spent 3 hours on a bed in an isolation cubicle. The interior of the cubicle was totally dark, and the only audible sound was the monotonous blower noise of an air-conditioner.

During the isolation period *S* could button press for visual, auditory, and kinesthetic stimulation. Three door bell buttons, distinguished by indentations, were mounted on the bed approximately 18 inches from *S*'s head. The buttons were connected to three indicator lights located in an adjacent *E*'s station.

Visual stimuli consisted of lighting the cubicle by a Penetray's motorized color wheel which was illuminated by a 150-watt reflector spotlight. The color wheel rotated amber, blue, green, and red at 4 revolutions per minute. An Amperite delay relay switch limited the length of visual stimulation to

¹Requests for reprints should be sent to John Hocking, who is now at the Kalamazoo Child Guidance Clinic, Kalamazoo, Michigan 49001.

15 seconds. Auditory stimulation consisted of four piano notes which were played on a tape recorder. Each time the auditory button was pressed, the tape lasted for 15 seconds or one revolution. A button press for kinesthetic stimulation permitted *S* freedom of movement both on and off the bed for 15 seconds. Since the bed springs emitted a squeaking noise whenever *S* moved, a record could be kept of movements during periods when the kinesthetic button was not pressed.

Procedure

After each *S* had been retested with the SSS and completed a brief autobiographical questionnaire, he was given the following instructions along with a demonstration of each form of stimulation.

You are asked to remain in the room for a period of time not more than 3 hours or until you voluntarily leave. You are asked to lie on the bed on your back. You are to stay awake. Please keep movements to a minimum and do not talk.

You will be able to get several forms of stimulation if you push a certain button. If you push the button with one mark on it, a light will go on for a period of time. If you push the button with two marks on it, tones will be played for a period of time. If you push the button in the middle position, you will be permitted to move from your position for a period of time. That is you may move around or get off the bed until a buzzer is sounded. When the buzzer sounds, please return to the supine position on the bed. The use of the buttons is entirely up to you. This experiment is not an endurance test or test of will power. People react in many ways and there is no correct pattern that we are looking for. The experimenter will remain next door in the observation room throughout the period of time. You will not be hurt in any way. Do you have any questions?

Analysis

The experimental period was divided into eighteen 10-minute units, and the button pressing responses for the three types of stimulation were totaled for each 10-minute period. The response variable was quite skewed in that most *Ss* made a low number of button presses, while a few *Ss* made a very high number. Consequently, nonparametric methods were used to analyze the data. The Mann-Whitney *U* test compared the high- and low-need *Ss* on the total number of button presses for each type of stimulation, as well as the total of all button pressing responses. Within each group, differences in type of stimulation requested were determined by a Friedman two-way analysis of variance.

RESULTS AND DISCUSSION

With the exception of visual stimulation, high scorers on the SSS demonstrated a greater need for stimulation than low scorers.

Table 1 shows that high-need *Ss* had a higher rate of total button pressing than the low-need *Ss*. The same was true of the subtotals for auditory and kinesthetic stimulation. However, the differences were not statistically significant. On the other hand, for visual stimulation low-need *Ss* had a significantly ($p = .05$) higher rate of button pressing than the high-need *Ss*.

Since a graphical representation of the responses of the two groups disclosed the greatest difference during the first hour, between-group comparisons were made both for total button presses and the subtotals for each type of stimulation during the first 60 minutes. Although the differences were in the direction noted above, they were not significant.

In terms of group preference for type of stimulation, high-need *Ss* showed a significant ($\chi^2 = 177.38$, $p < .001$) difference in favor of kinesthetic stimulation. For low-need *Ss* the difference was significant ($\chi^2 = 182.70$, $p < .001$) in favor of visual stimulation.

Although high scorers on the SSS had a greater rate of requesting stimulation than low scorers on the SSS, the failure to obtain a significant difference combined with the significant reversal on visual stimulation underscores the necessity of further scale validation. In further research four points should be considered. First, the effect of a longer period of sensory restriction, for example, 4-8 hours, should be assessed to determine whether the obtained direction of group differences is

TABLE 1
MANN-WHITNEY *U* TEST OF RATE OF BUTTON
PRESSING FOR HIGH AND LOW SCORERS
ON THE SSS

Response	High SSS		Low SSS		<i>U</i>
	Total rate	R_1^a	Total rate	R_2^a	
Visual	559	192	645	274	71.5*
Auditory	692	206	205	269	86
Kinesthetic	548	208	415	258	89
Combined	1,799	204	1,262	261	84

Note.—*N* = 15 in each group.

^a R = sum of ranks assigned to each group.

* $p = .05$.

maintained. Second, Zubek (1963) has suggested that exercise during isolation may provide enough variation of kinesthetic and proprioceptive stimulation to counteract some of the effects of restricted visual and auditory stimulation. The fact that the high-SSS group requested more kinesthetic stimulation than the low-SSS group and had a significantly greater preference for kinesthetic stimulation than for visual and auditory stimulation might be interpreted as an effort to lessen the unpleasant effects of restricted visual and auditory stimulation. Thus, a comparison of high and low scorers should perhaps be made under conditions in which kinesthetic stimulation is held to a minimum. Third, the high-SSS person is thought to be one who requires a variation in stimulation as well as amount of stimulation; consequently unless stimulation is quite varied and interesting, high-SSS persons are likely to consider it dull and to show less operant motivation.² Therefore, the stimuli provided in the experimental condition should have a variation dimension as well as an intensity dimension. Finally, additional analysis of the SSS items is needed. In this study, an item analysis of the SSS scores showed that 9 of the 26 items had low discriminatory power; that is, 70% or more of the Ss answered these items in the sensation seeking direction. On the other hand, 15 of the 26 items discriminated satisfactorily; that is, close to 50% of the Ss answered these items in the sensation seeking direction. On 12 of these items, Ss who answered in the sensation seeking direction button pressed substantially more or substantially less than Ss who answered the items

in the opposite direction. Thus, the relation between the SSS and overt behavior is at best an unclear one. Perhaps some clarity will be forthcoming from Zuckerman's (1967) efforts to replace the composite scale with specific factor scales that sample the stimulation dimensions separately.

REFERENCES

- ANSFIELD, P. A methodological study of isolation: Some effects upon activity and perception. *Dissertation Abstracts*, 1965, 26(2), 1179.
- BEXTON, W., HERON, W., & SCOTT, T. Effects of decreased variation in sensory environment. *Canadian Journal of Psychology*, 1954, 8, 70-76.
- DUFFY, E. *Activation and behavior*. New York: Wiley, 1962.
- FISKE, D., & MADDI, S. *Functions of varied experience*. Homewood, Ill.: Dorsey Press, 1961.
- GOLDSTEIN, K. Stimulus reinforcement during sensory deprivation. *Perceptual and Motor Skills*, 1965, 20, 757-762.
- HEBB, D. Drives and the central nervous system. *Psychological Review*, 1955, 62, 243-254.
- LEUBA, C. Toward some integration of learning theories: The concept of optimal stimulation. *Psychological Reports*, 1955, 1, 27-33.
- SCHULTZ, D. *Sensory restriction: Effects on behavior*. New York: Academic Press, 1965.
- VERNON, J., & MCGILL, T. Utilization of visual stimulation during sensory deprivation. *Perceptual and Motor Skills*, 1960, 11, 214.
- ZUBEK, J. P. Counteracting effects of physical exercises performed during prolonged perceptual deprivation. *Science*, 1963, 142, 504-506.
- ZUCKERMAN, M., KOLIN, E., PRICE, L., & ZOBB, I. Development of a sensation seeking scale. *Journal of Consulting Psychology*, 1964, 62, 1-20.
- ZUCKERMAN, M., & HABER, M. Need for stimulation as a source of stress response to perceptual isolation. *Journal of Abnormal and Social Psychology*, 1965, 70, 371-377.
- ZUCKERMAN, M. Progress report. December 1967, Pennsylvania Research Laboratories, Albert Einstein Medical Center.

² M. Zuckerman, personal communication, January 1968.

(Received July 18, 1968)

FOLLOW-UP OF THE LATER CAREERS AND LIVES OF 1,000 BOYS WHO DROPPED OUT OF HIGH SCHOOL¹

STARKE R. HATHAWAY,² PHYLLIS C. REYNOLDS,
AND ELIO D. MONACHESI

University of Minnesota

A dropout group and a nondropout group of high school boys were compared on several psychometric and demographic variables. The data both predated and postdated (by about 10 years) the time of dropout. For further analysis, the groups were subdivided by intelligence, additional training, and early versus late dropouts for the dropout group. The later lives of dropouts were characterized in general by relatively low socioeconomic status (SES), a downward social mobility, a higher incidence of criminal activity, and bigger families. Ninth-grade MMPI tests showed some personality differences between dropouts and nondropouts. When dropout and nondropout groups were matched by intelligence, most of the differences persisted. Training after dropout is of some help in future SES. Few differences were found between early and late dropouts.

During the 1950s somewhere between 15% and 30% of boys entering high school in the state of Minnesota did not graduate.³ Actual dropout rates are not known for Minnesota since the gathering of this kind of statistic presents difficulties such as accounting for school mobility, temporary dropouts, etc. In Lawrence, Kansas, over 27% of children who entered high school in 1950 did not graduate (Kansas State University, Bureau of Child Research, 1957); in Iowa, almost a quarter of adolescents who enter the ninth grade do not graduate (National Committee for Children and Youth, 1960).

During the 1960's, the Department of Labor estimates, some 7.5 million youngsters will drop out before high school graduation. About 2.5 million will not go beyond the eighth grade; two out of three will go not further than the tenth grade. Most will end their education before they have been exposed to the advice of the vocational counselor. Meanwhile, the level of training required for jobs is steadily rising. Each year there will be fewer openings, proportionately, for the unskilled worker. . . . Dismal as the dropout's immediate prospects are

when he leaves school, his long range outlook is even gloomier. The Department of Labor predicts that by 1970 white collar workers will constitute 45 per cent of the labor force, unskilled workers only 5 per cent [National Committee for Children and Youth, 1961, pp. 15 and 20].

These data attest to the critical social problem presented by the mass of adolescents who do not earn a high school diploma. The increasingly complicated technology in this country with its outcomes of automation and need for more and more skilled workers and persons who deal with theory and planning suggests that a high school diploma is becoming a minimum occupational requirement.

Although there have been many reports comparing dropouts to nondropouts (e.g., Cervantes, 1965; Hathaway & Monachesi, 1963; Livingston, 1958; Schreiber, 1967; Turchan, 1966) and reports on dropouts from the Midwest (e.g., Kansas State University, Bureau of Child Research, 1957; Nachman, Getson, & Odgers, 1964), surprisingly little has been published in the psychological, sociological, or educational literature about the later careers of the high school dropout. A recent report on occupational structure in America (Blau & Duncan, 1967) includes data on the later occupations of male dropouts, data which in many ways are far more extensive than the present; however, because of the nature of their study (retrospective), it could not provide data on some of the

¹ This research was supported by Research Grant 5 R01 MH 12187 from the National Institute of Mental Health and grants from the University of Minnesota.

² Requests for reprints should be sent to Starke R. Hathaway, Medical School, Division of Clinical Psychology, University of Minnesota, Minneapolis, Minnesota 55455.

³ Minnesota Department of Education, personal communication, April, 1968.

variables which are included in the present paper, for example, intelligence, trouble with the law, personality. The present study will be one of the few providing data on what happens to a rather large and representative sample of male high school dropouts, during approximately 10 years after the boys have left school. Although many questions are left unanswered, some evidence is provided on the relationship of dropout to data such as intelligence, marriage, size of family (of origin and of procreation), present occupation, training since high school, trouble with the law. The present authors also attempt a preliminary description of the personality factors associated with dropping out of school before completion.

METHOD

Samples

The dropout and control samples for this study were drawn from 11,329 Ss who comprised the statewide sample in a longitudinal study of adolescent personality and behavior which began in 1953 when the Ss were ninth-graders in Minnesota schools. For a more detailed description of the longitudinal study see Hathaway and Monachesi (1963); the sample was a random one from the state exclusive of the Minneapolis-St. Paul urban centers. It included the city of Duluth. It included 28% of the entire ninth-grade public school population of Minnesota in 1953-1954 and 35% of all ninth-graders outside the Minneapolis and St. Paul public school systems. It was impossible to record race when the data were gathered. Because of the racial composition of Minnesota, however, it is not far wrong to describe the sample as representative of predominately white, midwestern society.

The original number of boys in the sample was 5,701, of whom 1,002 (18%) dropped out of school before graduation (14% of the girls dropped out). Two follow-up studies were done on the entire 11,329; the first in 1957, shortly before high school completion, and the second over the years from 1962 to 1966. The mean age of the dropout boys was 27.9 years. Because two of the dropouts were lost in the second follow-up study, the total dropout *N* from which the table samples for this paper are drawn is 1,000. Except for the fact of high school graduation, the control sample boys were picked on a random basis within each school to equal the number of dropouts for the school.

Follow-Up

In the years 1953-1957, the initial and high school data were collected. Pertinent to this report were size of home community, socioeconomic status of parental

home, intelligence test scores, MMPI test results, and date of birth.

During the period 1962-1966, further follow-up data were gathered. The variables used in this report are date of marriage, present marital status, region of present residence, size of community, number of children, present occupation, type of training since high school, offenses against the law since 1957, emotional problems (as defined by known contact with a psychiatrist or psychologist in an in-patient or outpatient setting), status with the Armed Services.

Although 94% of the original statewide sample was contacted in some way during the 1962-1966 follow-up period, because of the usual difficulties inherent in a longitudinal study, the number of Ss that could be classified for each informational item used in this study varied from 50% to 89%.

RESULTS

This study summary of the dropouts is arbitrarily divided into five sections: (a) all dropouts compared to all controls; (b) dropouts compared to controls when grouped by high, medium, and low intelligence test scores; (c) dropouts who had some sort of further occupational training since high school compared to controls with further training; (d) dropouts with further occupational training since high school compared to dropouts with no further training; (e) boys who dropped out of school in the ninth and tenth grades compared to boys who dropped out in the eleventh and twelfth grades. Each of these will be considered separately. All analyses except one were done using the chi-square test. The authors felt justified in using only this test because of the nature of most of the data, which was a frequency count for each variable. Had we been able to do a follow-up in greater depth, we could have been able to often classify by interval scales, as did Blau and Duncan (1967) for socioeconomic status (SES). Unfortunately, this was not possible. The data presented here are derived from over 100 tables. Those included in the paper contain data that might be of more interest.*

All Dropouts Compared to All Controls

Differences with $\chi^2 < .001$ appeared in the following comparisons: age at the end of the

* Copies of all the tables can be obtained on request from S. R. Hathaway.

follow-up period—the dropout sample, although only .4 year older than the control sample in mean age, was negatively skewed with a median .5 year older; 69% were over the age of 27, whereas only 44% of the controls were over this age. Year of marriage—boys who drop out of school marry earlier, with the largest difference between them and the controls being the years when the controls were still in high school. Present marital status—although dropouts were more often divorced than controls, when this variable was controlled for date of marriage, there was no difference between dropouts and controls. Number of children—dropouts had larger families even when date of marriage was controlled.

Socioeconomic status of the family of origin (as measured by occupational level)—there was much fewer than the expected number of parents of dropouts who were in the professional, semiprofessional, or clerical, skilled categories (white collar) and much more than the expected number in the day laborer category. There were no significant differences between the number of dropouts and controls whose parents were in the semiskilled, slightly skilled (blue collar), farmer, or public assistance (unemployed) categories. Table 1 gives the percentages in each category. Classification was made in accordance with the Minnesota Scale for Paternal Occupations (Institute of Child Welfare, undated) both for this comparison and the one to follow. Present occupation of the dropouts appeared in far fewer than the expected numbers in the professional and semiprofes-

TABLE 1

DISTRIBUTION OF PARENT OCCUPATIONAL LEVELS
OF DROPOUTS COMPARED TO CONTROLS

Occupational level	Dropouts	Controls
Professional and semiprofessional	2	7
Clerical, skilled	15	20
Semiskilled	20	23
Slightly skilled	21	19
Day laborer	17	7
Farmer	25	23
Unemployed	1	1
Total	101	100

Note.—For dropouts, $N = 927$; for controls, $N = 975$. Distribution is in percentages. $\chi^2 = 81.39$, $df = 6$, $p < .01$.

TABLE 2

DISTRIBUTION OF SUBJECT OCCUPATIONAL LEVELS
OF DROPOUTS COMPARED TO CONTROLS

Occupational level	Dropouts	Controls
Professional and semiprofessional	2	16
Clerical, skilled	17	24
Semiskilled	26	21
Slightly skilled	30	16
Day laborer	11	3
Farmer	7	6
Unemployed	5	2
Student	3	11
Total	101	99

Note.—For dropouts, $N = 699$; for controls, $N = 816$. Distribution is in percentages. $\chi^2 = 204.06$, $df = 7$, $p < .001$.

sional category. This showed up to a lesser degree in the clerical-skilled category. In the semiskilled category no significant difference appeared between dropouts and controls, but in the categories of slightly skilled, day laborer, and unemployed there were many more than the expected number of dropouts. Because of the time of the follow-up study, a special category called student was provided. The proportion of controls who were still students was much greater than of dropouts. Table 2 gives the percentages of dropout Ss in each SES category compared to controls.

Social mobility—in an effort to further study the relationship of parental compared to S SES, the occupation of each S was compared to that of his family of origin. Dropout boys whose parents were in the professional, semiprofessional, and clerical, skilled combined categories (white collar) appeared in far less than the expected number in this category and in far more than the expected number in the semiskilled and slightly skilled combined categories. No differences were found between dropouts and controls from professional to clerical families who became day laborers or farmers or welfare recipients. Among those dropouts whose parents were in the semiskilled and slightly skilled combined categories there was less than the expected number who rose to the white collar level and more than the expected number who fell into the day laborer and public assistance category. Far fewer than the expected number of dropout boys whose parents were farmers were found in the white collar categories,

but there were no differences between dropouts and controls in the socioeconomic categories when the parents were farmers. A comparison of dropouts and controls whose parents were in the day laborer and unemployed combined categories showed no meaningful differences.

Type of training since high school—Table 3 shows that a comparatively large percentage of dropouts did not obtain any further training after leaving high school (72%); among those who did have additional schooling, there were far fewer than the expected number who had other than trade school training. There was no significant difference between the number of dropouts and controls who received a trade school training.

Offenses against the law—more dropouts than controls committed misdemeanors, felonies, and crimes severe enough to warrant a prison sentence, although there was no difference between the two groups when those committing only traffic offenses were compared. Emotional problems—although relatively few in actual number, the frequency of dropouts who were known to have received some form of psychiatric care since high school was significantly higher than the frequency of controls who had received psychiatric care.

Ordinal birth position—there was a tendency for dropouts to be the last born of large families (≥ 7 children in the family) and a slight tendency for them to be firstborn of large families. This is in some contrast to the findings of Blau and Duncan (1967) that oldest and youngest children are better educated than their middle siblings. Due to the nature of the present analysis, however, we

would concur with Blau and Duncan that the dropout tends to come from a large family. Although the present data cannot show this, other studies suggest that older brothers' and sisters' dropping out of school is an influence on *S* to drop out (National Committee for Children and Youth, 1961), as is having a mass of younger siblings to support (Blau & Duncan, 1967).

Personality—more than the expected number of dropouts produced invalid ($L \geq 10$, $F \geq 22$, $P \geq 40$) MMPI profiles in the ninth grade. Some but not all of this invalidity is interpretable as indicating poor reading ability or noncooperativeness in the testing situation. More than the expected number of dropouts also had MMPI profiles with the highest scale being 8 (Schizophrenia), suggesting that there was a tendency for dropouts to be negative, difficult, odd, apathetic, or lacking in social graces. Less than the expected number had profiles where the highest scale was 3 (Hysteria) or 9 (Hypomania), suggesting a tendency for dropouts to be less characterized by such qualities as being idealistic, articulate, socially conforming, optimistic, decisive, expansive. Table 4 gives the chi-square computation on the numbers of dropouts and controls for each MMPI high point.

When the mean scores on each scale were compared, however, a somewhat different picture appeared. Table 5 gives the means and standard deviations for the validity scales and the 10 clinical scales of the MMPI for those boys with valid profiles. Because of the large numbers in the samples, some small differences between the means were highly significant; yet the mean scores on scales *F* (validity), 4 (Psychopathic Deviate), and 8 (Schizophrenia) for the dropouts seem meaningfully higher than those of the controls. A high *F* scale can signify an unusual amount of psychological complaint; a high Scale 4 suggests disregard for rules and feelings of rebelliousness; and a high Scale 8, as mentioned, suggests feelings of isolation, negativism, and apathy. Whereas the chi-square results of Table 4 are based on the first high point of each boy's MMPI profile, the mean scale values of Table 5 give an idea of the actual average elevation of each

TABLE 3

PERCENTAGES OF DROPOUTS AND CONTROLS WITH FURTHER TRAINING AFTER HIGH SCHOOL

Further training	Dropouts	Controls
Professional, education, business, etc.	8	42
Trade school	20	16
No further training	72	41
Total	100	99

Note.—For dropouts, $N = 820$; for controls, $N = 888$. $\chi^2 = 270.56$, $df = 2$, $p < .001$.

TABLE 4
MMPI HIGH POINT COMPARISONS
FOR DROPOUTS AND CONTROLS

Scale	Dropouts			Controls			Total	χ^2
	o	χ^2	e	o	χ^2	e		
1	12	.29	14	16	.29	14	28	.58
2	33	.89	28	23	.89	28	56	1.78
3	7	4.27	15	22	4.27	14	29	8.84
4	160	.42	152	138	.44	146	298	.86
5	16	1.64	22	28	1.64	22	44	3.28
6	38	.82	44	48	.86	42	86	1.68
7	45	.19	48	49	.20	46	94	.39
8	255	6.65	217	171	6.91	209	426	13.56
9	118	9.26	156	189	9.56	151	307	18.82
0	38	2.88	50	61	2.94	49	99	5.82
Invalid	110	6.70	86	59	6.94	83	169	13.64
Total	832	34.01		804	35.24			

Note.— $\chi^2 = 69.25$, $df = 10$, $p < .001$. o = actual number; e = estimated number.

of the scales. Dropouts clearly tended to get more elevated MMPI profiles than did controls.

The following comparisons yielded little or no statistically significant differences: region of present residence—when dropouts are compared to controls on remaining in Minnesota versus moving elsewhere, or by more extensive breakdown of region of residence (e.g., Minnesota, California, Pacific Northwest, Midwest, New England, South, etc.), there was no significant difference between

TABLE 5

MEANS AND STANDARD DEVIATIONS OF THE MMPI
SCALES FOR DROPOUTS AND CONTROLS:
VALID PROFILES ONLY

Scale	Dropouts		Controls		t
	M	SD	M	SD	
L	3.89	1.96	3.84	2.06	.54
F	8.57	4.90	6.18	4.17	11.27*
K	12.93	4.45	13.69	4.67	3.58*
1 (+.5K)	52.52	10.31	51.52	9.00	2.20
2	54.67	10.81	51.69	10.45	5.98*
3	52.55	8.42	51.91	7.86	1.68
4 (+.4K)	62.90	11.09	58.44	9.84	9.06*
5	50.21	9.01	51.24	8.72	2.40
6	58.33	11.31	56.00	9.92	4.67*
7 (+K)	60.58	10.91	58.25	9.95	5.18*
8 (+K)	66.50	14.85	61.87	12.90	7.09*
9 (+.2K)	61.17	11.30	58.84	11.40	4.38*
0	55.42	7.77	54.07	8.22	3.61*

Note.—For dropouts, $N = 884$; for controls, $N = 941$. Dropout mean profile '849760—4:9:13. Control mean profile '89476—4:6:14.
* $p < .002$.

the two groups. Community size—dropouts did not tend to congregate in any one size of community, city, large town, small town, or farm. Armed Services—branch of service did not generally distinguish dropouts from controls. Although the chi-square on status with the service was significant beyond $p = .005$, the only cells that meaningfully distinguished dropouts from controls were those for the National Guard. There were fewer dropouts than controls in the Guard. Dropouts were not discharged from the Armed Services for any reason with higher frequency than controls.

Dropouts Compared to Controls when Grouped by Intelligence

Percentile scores from an eleventh or ninth grade American College Entrance (ACE) exam or, in a few cases, percentiles derived from an IQ test were used for these analyses. Grouping was as follows: high 69–100 percentile; medium 40–68 percentile; low 1–39 percentile.

Comparison on the following variables were significant at $p < .005$: present marital status—although there was no difference among the three dropout groups themselves, when high intelligence dropouts were compared to high intelligence controls, there was a tendency for more than the expected number of dropouts to be separated or divorced. Year of marriage—again the three dropout groups did not differ, but all three differed significantly from their comparable control group, each group marrying earlier than the controls.

Present occupation—the overall table comparing the three dropout groups was significant at $p < .001$; however, it was primarily the medium intelligence dropout group that showed significant differences. It was this group, *not* those of high intelligence, that had more than the expected number in the professional, semiprofessional, and clerical, skilled categories. Even when the medium intelligence dropouts were compared to the medium intelligence controls there was no difference between the two groups in the clerical, skilled category. Both the high and low intelligence dropout groups differed greatly from their comparable control groups, and both in the same categories: fewer than the expected

number in the white collar category and more in the slightly skilled, day laborer, and unemployed categories.

Type of training since high school—chi-square values on this variable were all significant at $p < .001$. Among the dropout groups themselves there is a pattern of more than expected high intelligence boys who received training in the academic or business school lines, more than the expected number of medium intelligence boys who went to trade school, and more low intelligence boys who had no further training. The most marked difference when these groups were compared to their controls appeared in the relatively small number who pursued academic or business lines among the dropouts, no matter what the intelligence. For the boys in the dropout or the control groups who went to trade school there were no significant differences, that is, a high intelligence dropout was just as likely to go to trade school as a high intelligence boy with a high school diploma.

Offenses against the law—no differences appeared when the three dropout groups were compared; when each was compared to its control, however, it was only between low intelligence dropouts and low intelligence controls that there were no differences. Both high and medium intelligence dropouts had significantly more offenses ranging from misdemeanors to prison sentences than did their control groups.

MMPI high point—when just those high points that distinguished dropouts from controls were considered (Scales 3, 8, 9), only those dropouts who scored high on intelligence tests got more than the expected number of profiles with 8 as the highest scale. Intelligence did not differentiate on either Scale 3 or 9.

On some of the variables there were no significant differences between dropouts and controls when grouped by intelligence test scores: region of present residence showed no differences except between the medium intelligence groups. More dropouts had moved away from Minnesota. Community size and number of children showed no meaningful differences. Status with the Armed Services developed a difference only between the low intelligence groups, which had less than the

expected number of dropouts in the National Guard.

When the dropouts were grouped by reason for dropping out of school, that is, failing or lack of interest, work, institutionalized, Armed Service, there were no significant frequency differences among the high, medium, or low intelligence boys.

Dropouts with Further Training Compared to Controls with Further Training

For the grouping in these analyses, dropouts and controls were included only if they were known to have had training of any kind after high school. Training ranged from advanced degrees to trade school. The purpose for study of these groups was to explore the discrimination that relates to a high school diploma. The following chi-squares were significant ($p < .001$).

Present marital status—the dropout group was less likely to be single and more likely to be divorced than the control group.

Type of training since high school—far greater than the expected number of dropouts went on to trade school training rather than to pursue a more academic avenue; but when present occupation was evaluated, there was no difference between controls and dropouts in the clerical, skilled category. Table 6 shows that when compared to controls, dropouts congregated among the semiskilled and slightly skilled and day laborer categories despite their training.

Offenses against the law—the group of

TABLE 6
COMPARISON OF PRESENT OCCUPATION BETWEEN
DROPOUTS AND CONTROLS WITH TRAINING
AFTER HIGH SCHOOL

Occupation	Dropouts	Controls
Professional and semiprofessional	6	27
Clerical, skilled	28	26
Semiskilled	29	18
Slightly skilled	21	6
Day laborer	5	0
Farmer	3	2
Unemployed	2	2
Student	6	18
Total	100	99

Note.—For dropouts, $N = 189$; for controls, $N = 454$. Comparison is in percentages. $\chi^2 = 90.51$, $df = 7$, $p < .001$.

dropouts who had further training, like the total dropout group, got into significantly more trouble with the law than did the controls. More of them than controls committed misdemeanors, felonies, and crimes severe enough to warrant a prison sentence.

Comparisons that showed little or no differences were year of marriage, number of children, region of present residence, size of community, and MMPI high points.

Dropouts with and without Training after Quitting High School

An attempt was made to find demographic signs of future success despite high school dropout, that is, even if a boy quits school, are there any indications that he makes up for this deficit in some way, and can one tell which boys are more likely to make up for the deficit? Again, training is defined as anything ranging from return to high school plus continuation on to a professional degree to trade school.

Boys who obtained further training showed a tendency ($p < .05$) to come from families that were professional, semiprofessional, or semiskilled and to not come from farm families. Although less than expected of them came from a city or suburb, far greater than the expected number grew up in large towns. Relatively few came from farms whether or not their fathers were actually farmers. While or after they had their training, however, they tended to move to the city, and even if they originally came from a farm, they did not stay or return there. Also they moved away from Minnesota ($p < .001$) with much higher frequency than dropouts who had no training.

Intelligence showed some association with continued training. Fewer than the expected number of low intelligence dropouts went on for training, while clearly more than the expected number of both medium and high intelligence boys got more training.

Whether the family of origin was a broken one by death or divorce seemed to make no difference in determining that a dropout went back to school. And between the two groups themselves there was no difference in rate of divorce, although those dropouts who got no further training tended to marry earlier.

There was a distinction between the groups on present occupation. As Table 7 shows, the percentage of boys with further training who were in the white collar category was more than twice that of boys without training. As would be expected, the nontrained boys congregated in the low SES categories, and only 1% of them were students at the time of follow-up.

The ninth-grade MMPI testing distinguished weakly at the $p < .10$ level between dropouts who did and did not get further training. The MMPI high point pattern was similar to that produced when all dropouts were compared to all controls: boys who went back to school less often had a high 8, more often had a high 9. They also less often had invalid profiles, and although it did not appear on the overall comparisons, they more often had a high point of 5 (Masculinity-Femininity), a scale that is frequently associated with educational persistence (Dahlstrom & Welsh, 1960).

Early (Ninth and Tenth Grade) Dropouts Compared to Later (Eleventh and Twelfth Grade) Dropouts

There were few differences in the later lives of early dropouts compared to late dropouts. There were no significant differences on year of marriage, marital status, present area of residence, offenses against the law, later emotional problems, or status with the Armed Services.

TABLE 7

DROPOUTS WITH FURTHER TRAINING AFTER HIGH SCHOOL COMPARED TO DROPOUTS WITH NO TRAINING, ON PRESENT OCCUPATION

Occupation	No training	Training
Professional, semiprofessional, clerical-skilled	13	34
Semiskilled	25	29
Slightly skilled	34	21
Day laborer	11	5
Farmer	9	3
Unemployed	6	2
Student	1	6
Total	99	100

Note.—For dropouts with no training, $N = 470$; for dropouts with training, $N = 189$. Comparison is in percentages. $\chi^2 = 71.14$, $df = 6$, $p < .001$.

Early dropouts tended to live on farms more than did later dropouts. Also early dropouts originally came from farms with a much higher frequency ($p < .001$).

When these groups were compared on type of training after leaving high school, less than the expected number of early dropouts pursued academic lines; not surprisingly it was the later dropouts who more frequently returned to complete high school and from that continued in regular academic lines or attended some sort of business school ($p < .005$).

DISCUSSION

The most marked feature that distinguishes the male high school dropout from those boys who receive their high school diploma is the occupational level achieved by the time these men are in their late 20s. To a large degree the dropout of the 1950s is the slightly skilled or unskilled laborer of the 1960s. Even if a bright boy drops out of school, the data show that he is much less likely than his bright schoolmate who graduates to get later training of any kind and that he is much more likely to subsequently fall into the ranks of the semi-skilled, slightly skilled, day laborer, and unemployed. In view of the nationwide search for talent, this finding is important: boys of high ability *are* being lost.

Downward social mobility of dropouts is another finding of this study. The data show that boys who do not finish high school tend to drop in socioeconomic level from their family of origin regardless of what level their parents were. Social class does determine to some extent whether or not a boy finishes high school (88% of boys from professional, semiprofessional, and clerical, skilled families finished high school compared to 78% of boys from semiskilled, slightly skilled, and day laborer families), but once a boy drops out of school, the social class of his parents has little relationship to his own social class by the time he is in his late 20s. These findings do not exactly agree with those of Livingston (1958), who found that occupational status of the principal wage earner in the family of origin was an insignificant factor in whether or not a boy dropped out of school, but they do agree with the findings of a much larger

study (Blau & Duncan, 1967), where it is shown that the number of years of schooling completed has a much higher relationship with later career than does the father's occupational status.

Until drastic changes are made in primary and secondary school curricula in the United States (cf. Schreiber, 1967, for discussions of possibilities), it can be assumed that there will always be a certain proportion of boys who will drop out of high school. At age 16 they can do so legally in Minnesota and many other states. If one had to select a group to drop out, perhaps the most justifiable selection would be made by using intelligence as a criterion, that is, boys with a below-average intelligence would appear to have the most difficulty in keeping up with the "academic" types of curricula presented by the usual high school of today. A fair question to ask is how much of a social problem would boys of below-average intelligence who drop out of school become. The data in this study show that although there is a tendency for low intelligence dropouts to be separated and divorced more often than high school graduates, they are not separated or divorced with any higher frequency than dropouts of average or above average intelligence. Although low intelligence dropouts tend to congregate in the low SES categories when compared to low intelligence controls, they are not actually unemployed or on public assistance to any greater extent than the controls, and when compared to average and above-average intelligence dropouts, there is essentially no difference between the ultimate SES level. They do not have larger families than either the controls with a below-average intelligence or the higher intelligence dropouts, thus it cannot be said that they are potential higher producers of future social problems than are other dropouts. Even when trouble with the law is considered, the below-average intelligence dropout is no different from his control *S* or from the average to above-average intelligence dropout.

This sort of reasoning makes it clearer that attempting to limit high school dropouts to just those boys who are the most "eligible" candidates for not continuing, that is, those who are of too limited intelligence to benefit

from academic curricula, would do little if anything to solve the social problems presented by high school dropouts. The culprit may be the value placed on the high school diploma. Without it a boy, no matter what his basic intelligence, tends to end up in low-paying, unsteady jobs or perhaps unemployment as he grows older.

As was stated earlier, boys of talent, if talent is broadly defined by above-average intelligence, *are* being lost. How severe is this loss? In the first place, 14% of the dropouts have intelligence test scores at or above the 69th percentile when compared to their age group in high school. These boys are no more mobile than nondropouts, that is, they remain in the state to the same degree and they do not congregate in any one community size area more than nondropouts. Sixty-nine percent of them stay in Minnesota, yet 59% receive no further training after quitting school, and only 15% reach the occupation level of clerical, skilled or above (71% fill the ranks of semiskilled, slightly skilled, day laborer, and unemployed). As adults they get into much more serious trouble with the law than does the group of boys who graduate from high school (37% of high intelligence dropouts have committed offenses ranging from misdemeanors to ones severe enough to warrant a prison sentence as compared to 9% of the high school graduates). Clearly, the loss seems rather severe: these boys tend to become a burden to the state; their talent seems to be spent in fighting the laws rather than making them. These boys, of all of the dropouts, have the intelligence potential to stay in school at least until high school graduation. With a high school diploma they could have at least qualified for training programs for skilled workers that so often now require the diploma for entrance. A study by Gibson, Higgins, and Mitchell (1967) offers a note of pessimism on preserving the talent of dropouts. Of 6,008 dropouts in Toledo (both male and female), 7% had potential for college and were within commuting distance to a college. After individual contact, only 3% (13 persons) of this group entered college, and of these 13, only 11 persons completed the freshman year, and only 4 of these were making satisfactory progress.

When a comparison is made between dropouts who do and do not get further training after quitting school, those that do go on can be characterized by the following features: they are more likely to come from families at a professional or semiprofessional SES level and, interestingly, from semiskilled families. Since they are also more likely to profit by their training and secure white collar or high blue collar jobs, the downward social mobility trend does not show up here as it does when dropouts in general are compared to high school graduates. They are more likely to be of average or above-average intelligence. Although they do not grow up in the cities and suburbs, they tend to migrate there as adults, and they tend to leave Minnesota.

Dropouts in general become burdens to the state; the state loses those dropouts most likely to succeed (36% of them go to the West Coast; 23% to California alone). Whereas intelligence and SES do to some degree seem to be determinants in returning to school for training, the factor of coming from intact or broken homes is not. It is the boy who drops out early (ninth or tenth grade), who comes from a farming family or an unskilled labor family in the city, and who does not have good intellectual potential that not only drops out of high school but never seeks further training of any kind.

Reports of personality testing on dropouts compared to high school graduates have been infrequent. Cervantes (1965) administered the TAT to a group of known dropouts and controls in New Orleans and Boston (the exact number in each group is not clear from the report, but it is possibly 50 dropouts and 50 matched controls), and from these protocols he arrived, in an unreported way, at a list of psychological tendencies of the dropout. This list includes 36 descriptive items, including troubled, hostile, antagonistic, instinctoid [sic], unconventional, maladjusted, nonroutinized. In contrast to Cervantes' data, the present report gives results on personality testing *before* the time of drop out and attempts to show personality characteristics that can be used in a predictive sense.

The ninth-grade personality testing with the MMPI yields distinctions between later

dropouts and nondropouts. The results in this paper differ somewhat from those presented earlier (Hathaway & Monachesi, 1963) because they include MMPI profiles considered as invalid earlier; for this paper an F score less than 22 was considered as valid. There is some association between MMPI high point in the ninth-grade and later school dropout ($\phi = .20$, $N = 1,636$), but the predictive value of knowing a high point is only slight ($\lambda B = .17$). This does not entirely preclude a discussion of the dropout personality types, however. On the whole, dropouts tend to get a more elevated profile over all scales except Scale 5. Especially elevated are Scales 8 and 4. More dropouts also tend to have profiles where 8 is the highest point in the profile (31% of all the dropouts have 8 as the highest point, compared to 21% of the controls). An 8 as the highest point occurs more among high intelligence dropouts compared to high intelligence controls and among dropouts who do not get further training compared to dropouts who do. Whereas the elevated Scale 4 suggests that future dropouts harbor feelings of rebelliousness and lack feelings of social conformity in the ninth grade, the more frequent occurrence of 8 as the highest point suggests early feelings of apathy, difficulty in thinking straight, and tendencies to withdraw. The high school counselor should be on the alert when he finds a boy with a high 8 MMPI profile; he could be a future dropout, even if he is bright. And if he does drop out of school, he is likely to become a real social dependent.

The Scale 9 findings suggest that a little elevation of 9 is a good thing but too much is detrimental. Dropouts have more elevated 9s but high school graduates more frequently have 9 as the highest point on the profile. An elevated 9 is associated with a high energy level; perhaps those dropouts with elevated 9s had too much energy to stick to school. Among the statewide sample as a whole, a high 9 profile was the most common (Hathaway & Monachesi, 1963); these adolescent boys in general are more often characterized by adjectives such as sociable, enthusiastic, forward, talkative, and verbal; and boys who graduate from high school seem to be especially characterized by such adjectives. The

further finding that there are more than the expected number of high point 3 profiles among those of the controls is in line with the above description. High point 3s are described as socially conforming, persevering, energetic, cooperative.

CONCLUSION

Although the picture painted by this study is dark, it is not quite so black as that from studies that cover large urban areas or the South. In the late 1950s, at least, Minnesota had a comparatively low dropout rate for high school boys—18% versus a percentage like 59 for a slum area of a large city (Conant, 1961). Only 5% of the Minnesota dropouts were unemployed by their late 20s (which is only slightly more than the national figure of 3.8% for white males aged 25–34 in 1962, Bienstock, 1967), although approximately 50% of them are likely candidates for future unemployment if the Department of Labor predictions are correct. The deep problem in the country today is concerned with the futures of deprived persons, especially members of racial minority groups, where even a high school diploma is of little value in securing employment. The data of this study suggest that in predominately white, mid-western America, the high school diploma is a passport to occupational security. They show, however, that certain background characteristics of adolescent boys are influential in determining whether or not a diploma is attained: socioeconomic status of the parents, family size, intelligence, and personality. These characteristics have long been recognized and discussed in the literature, and so the findings in this report do not come as a surprise. What is surprising are the findings that if a boy does drop out of high school, these same characteristics have little influence on his later career.

REFERENCES

- BIENSTOCK, H. Realities of the job market for the high school dropout. In D. Schreiber (Ed.), *Profile of the school dropout*. New York: Random House, 1967.
- BLAU, P., & DUNCAN, O. D. *The American occupational structure*. New York: Wiley, 1967.

- CERVANTES, L. F. *The dropout: Causes and cures*. Ann Arbor: University of Michigan Press, 1965.
- CONANT, J. B. *Slums and suburbs: A commentary on schools in the metropolitan areas*. New York: McGraw-Hill, 1961.
- DAHLSTROM, W. G., & WELSH, G. S. *An MMPI handbook*. Minneapolis: University of Minnesota Press, 1960.
- GIBSON, R. L., HIGGINS, R., & MITCHELL, M. H. The high school dropout goes to college. *Personnel and Guidance Journal*, 1967, 45, 824-827.
- HATHAWAY, S. R., & MONCHESI, E. D. *Adolescent personality and behavior*. Minneapolis: University of Minnesota Press, 1963.
- INSTITUTE OF CHILD WELFARE, UNIVERSITY OF MINNESOTA. *Minnesota Scale for Paternal Occupation*. Minneapolis: Author, undated.
- KANSAS STATE UNIVERSITY, BUREAU OF CHILD RESEARCH. *Children and Youth in Kansas*. Manhattan: Author, 1957.
- LIVINGSTON, A. H. High school graduates and dropouts: A new look at a persistent problem. *School Review*, 1958, 66, 195-203.
- NACHMAN, L. R., GETSON, R. F., & ODGERS, J. G. *Ohio study of high school dropouts: 1962-1963*. Columbus: Division of Research and Division of Guidance and Testing, Ohio State Department of Education, 1964.
- NATIONAL COMMITTEE FOR CHILDREN AND YOUTH, 1960 WHITE HOUSE CONFERENCE. *The states report on children and youth*. Washington, D. C.: United States Government Printing Office, 1960.
- NATIONAL COMMITTEE FOR CHILDREN AND YOUTH. *Social dynamite: The report of the conference on unemployed, out-of-school youth in urban areas*. Washington, D. C.: United States Government Printing Office, 1961.
- SCHREIBER, D. (Ed.) *The school dropout*. Washington, D. C.: National Education Association, 1964.
- SCHREIBER, D. (Ed.) *Profile of the school dropout*. New York: Random House, 1967.
- TURCHAN, D. G. A comparison of high school dropouts matched with high school persisters in an urban school system. *Dissertation Abstracts*, 1966, 27(2-A), 335-336.

(Received July 26, 1968)

BRIEF REPORTS

DRIVER ACCIDENTS AND THE NEUROPSYCHIATRIC PATIENT¹

MATTHEW W. BUTTIGLIERI

Veterans Administration Hospital, Sepulveda, California

M. I. CHARLES E. WOODSON, MARIE GUENETTE, AND MAE THOMSON

*Institute of Transportation and Traffic Engineering
University of California, Los Angeles*

In one of the most sophisticated studies of accident-related personality traits, Conger, Gaskill, Glad, Hassel, Rainey, and Sawrey (1964) hypothesized 13 variables to be significantly related to accident frequency. These variables, for example, amount of underlying hostility, reality orientation, conventionality, extent of fantasy preoccupation, ability to tolerate various tensions, etc., as well as the hypothetical composite accident type described by Conger et al., contain many personality characteristics that one would expect to find in emphasized or exaggerated form in the hospitalized mental patient.

To explore the degree of relationship between driver records of neuropsychiatric patients (NP) and their scores on psychological tests administered as part of the routine admission procedure, and to develop criteria for use in the identification of high-risk drivers among neuropsychiatric patients, the driving records of 533 NP patients admitted to Sepulveda Veterans Administration hospital from August 1964 through July 1967 were compared with published data for a random sample ($N = 86,717$) of California drivers. Their accident records did not differ statistically, but

violation records of the NP sample were higher ($p < .001$) than the compared group.

A regression analysis approach was used to examine the possibility that variables from the MMPI, Wonderlic Personnel Test, and selected biographical variables would aid in the prediction of the patients' accident record. Three variables which were significantly related to accident records were selected using the first 2-year sample; Schizophrenic Reaction, other; Psychoneurotic diagnosis; and *Ma* score. A drink (yes/no) variable was added on rational grounds. Diagnosis Schizophrenic Reaction, other; *Ma* score; and drinking were validated by the third-year sample. The resulting regression equation accounted for 9% of the variance in accident involvement.

The possibility of improving prediction of accident records by selection of subpopulations in which prediction is more effective was also explored. Lack of comparative data on the non-hospitalized population prevented comparison, but the results are suggestive for further studies.

In general, the results indicated that patients with diagnosis Schizophrenic Reaction, other, had a lower accident rate, while patients who drank or had a high *Ma* score had a higher accident rate.

REFERENCES

- CONGER, J. J., GASKILL, H. S., GLAD, D. D., HASSEL, L., RAINY, R. V., & SAWREY, W. L. Psychological and psychophysiological factors in motor vehicle accidents. In W. Haddon et al. (Eds.), *Accident research*. New York: Harper & Row, 1964.
- STATE OF CALIFORNIA, DEPARTMENT OF MOTOR VEHICLES. Driver record by age, sex and marital status. June 1965, 24-25.

(Received September 24, 1968)

¹ An extended report of this study may be obtained without charge from Matthew W. Buttiglieri, Assistant Chief, Psychology Service, Veterans Administration Hospital, Sepulveda, California 91343, or for a fee from the American Society for Information Science. Order NAPS Document 00364 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001; remitting \$1.00 for microfiche or \$3.00 for photocopies. Requests for reprints should be sent to the first author at the above address.

PERCEIVED CHILD-REARING PRACTICES AND SELF-DISCLOSURE PATTERNS¹

JOSEPH A. DOSTER AND BONNIE RUTH STRICKLAND

Emory University

In a time of supposed alienation of youth from their elders, an investigation of the relationship between parental reinforcement and patterns of verbal expression becomes of particular importance. Jourard (1964) describes self-disclosing behavior and reports females to be more disclosing than males; mothers disclosed to more often than fathers or friends; and differences in what people reveal about themselves. Heilbrun (1964), using the Parent-Child Interaction Rating Scale (P-C), found mothers receiving higher nurturance ratings than fathers from their children and this difference was greater for male Ss than females.

This study attempts to relate perceived child-rearing practices and self-disclosure as well as to replicate some of Jourard's findings. Hypotheses are (a) home environments seen as nurturant will yield greater disclosure than low-nurturant homes and (b) parental nurturance will have greater influence on disclosure to parents than to friends. College Ss were given the Self-Disclosure Questionnaire (SDQ); the P-C; and the Marlowe-Crowne Social Desirability Scale (M-C SDS—Crowne & Marlowe, 1964).

A significant positive relationship was found between nurturance ratings for each parent ($r = .61$, $p < .01$, $N = 110$), suggesting that offspring see similarity in the child-rearing attitudes of their parents. Analysis of variance ($N = 80$) revealed a significant main effect of greater overall disclosure by Ss who saw parents as highly

nurturant than by Ss who saw their family background as low in nurturance. A significant interaction of Parental Nurturance \times Target Person indicates that Ss from low-nurturant homes disclose more to friends than to parents, with the reverse being the case for Ss perceiving their parents as nurturant.

Significant main effects for sex differences and for differences in disclosure output to target persons were not found. However, trends are toward more disclosure to mother than to father, and more disclosing on the part of females than males. A significant Sex \times Target Person interaction showed that females disclose more to mother and female friend than to father and male friend. Males are about equally disclosing to mother, father, and male friend, but less to female friend. The Ss reported revealing significantly more about public aspects of self than private or personal aspects and friends are more often the recipients of public information, and parents the recipients of personal information. Intercorrelations of the M-C SDS with SDQ and P-C were generally low and nonsignificant. Thus, a social desirability response set does not appear to be operating in the data.

Surprisingly, our student sample indicates a turning to parents for personal disclosure and perhaps not so surprisingly a relationship between stated self-expression to parents and perceived parental nurturance. Focusing attention on the nature of messages communicated is probably the next logical step in research.

REFERENCES

- CROWNE, D., & MARLOWE, D. *The approval motive: Studies in evaluative dependence*. New York: Wiley, 1964.
- HEILBRUN, A. Parent model attributes, nurturant reinforcement and consistency of behavior in adolescents. *Child Development*, 1964, 35, 151-167.
- JOURARD, S. *The transparent self*. New York: Van Nostrand, 1964.

(Received October 28, 1968)

¹ An extended report of this study may be obtained without charge from Joseph A. Doster, Department of Psychology, Emory University, Atlanta, Georgia 30322, or for a fee from the American Society for Information Science. Order NAPS Document 00365 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York 10001; remitting \$1.00 for microfiche or \$3.00 for photocopies. Requests for reprints should be sent to Bonnie Strickland, Department of Psychology, Emory University, Atlanta, Georgia, 30322.

PSYCHIATRY

Journal for the Study of Interpersonal Processes

Over Thirty Years of Distinguished Interdisciplinary Communication

Selected Articles in Vol. 32, 1969

- Albert Cain.* Special Isolated Abilities in Severely Psychotic Young Children.
William Caudill, Helen Weinstein. Maternal Care and Infant Behavior in Japan and America.
Joel Fischer. Negroes, Whites, and Rates of Mental Illness: Reconsideration of a Myth.
David Forrest. New Words and Neologisms. With a Thesaurus of Coinages by a Schizophrenic Savant.
Neil Friedman. Africa and the Afro-American: The Changing Negro Identity.
Erving Goffman. The Insanity of Place.
John Hoskin et al. A High Incidence of Suicide in a Preliterate-Primitive Society.
Francis King. Marihuana and LSD Usage Among Male College Students.
Theodore Lidz. The Influence of Family Studies on the Treatment of Schizophrenia.
Norman Polansky, Elizabeth Harkins. Psychodrama as an Element in Hospital Treatment.
Richard Sternbach. Clinical Responsibility and the Psychologist.
Edwin Weinstein, Olga Lysterly. Symbolic Aspects of Presidential Assassination.
Naomi Wenner et al. Emotional Problems in Pregnancy.
L. S. Zegans et al. A Mental Health Center's Response to Racial Crisis in an Urban High School.

BONUS TO ALL 1969 SUBSCRIBERS

30-YEAR INDEX, 1938-1967

(\$10 to nonsubscribers)

Subscriptions: \$10.00, foreign \$10.80; payable in advance.

Published in February, May, August, and November.

Order from: Dept. A, PSYCHIATRY

1610 New Hampshire Ave., N.W., Washington, D. C. 20009

Journals of the AMERICAN PSYCHOLOGICAL ASSOCIATION

Subscription Rates for Calendar Year 1969

AMERICAN PSYCHOLOGIST. Official papers of the Association and articles on Psychology. Monthly. The 1969 volume is Vol. 24. Subscription: \$10.00 (foreign \$10.50). Single copy \$1.00.

CONTEMPORARY PSYCHOLOGY. Critical reviews of books, films, and research material in the field of psychology. Monthly. The 1969 volume is Vol. 14. Subscription: \$10.00 (foreign \$10.50). Single copy \$1.00.

JOURNAL OF ABNORMAL PSYCHOLOGY. Theoretical and research conditions in behavior pathology. Bimonthly; first issue appears in February. The 1969 volume is Vol. 74. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00. (Continues volume sequence of the former Journal of Abnormal and Social Psychology.)

JOURNAL OF APPLIED PSYCHOLOGY. Applications of psychology in business and industry. Bimonthly; first issue appears in February. The 1969 volume is Vol. 53. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF COMPARATIVE AND PHYSIOLOGICAL PSYCHOLOGY. Original contribution in the field of comparative and physiological psychology. Beginning in 1969. Monthly. Three volumes per year, four issues in each volume. The 1969 volumes are Vols. 67, 68 and 69. Subscription: \$30.00 for 3 vols. (foreign \$30.50). Single copy \$3.00.

JOURNAL OF COUNSELING PSYCHOLOGY. A primary publication medium for research in counseling theory and practice. Bimonthly; first issue appears in January. The 1969 volume is Vol. 16. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF CONSULTING AND CLINICAL PSYCHOLOGY. Research in clinical psychology: psychological diagnoses, psychotherapy, personality, psychopathology. Bimonthly; first issue appears in February. The 1969 volume is Vol. 33. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF EDUCATIONAL PSYCHOLOGY. Studies of learning and teaching; measurement of psychological development, psychology of school subjects, methods of instruction, school adjustment. Bimonthly; first issue appears in February. The 1969 volume is Vol. 60. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY. Original contributions of an experimental character. Monthly. Beginning 1969: Four volumes per year, three issues in each volume. The 1969 volumes are Vols. 79, 80, 81 and 82. Subscription: \$40.00 for 4 vols. (foreign \$40.50). Single copy \$4.00.

JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY. Theoretical and research papers on personality dynamics, group process, and the psychological aspects of social structure. Monthly. Three volumes per year, four issues in each volume. The 1969 volumes are Vols. 11, 12 and 13. Subscription: \$30.00 for 3 vols. (foreign \$30.50). Single copy \$3.00.

PSYCHOLOGICAL ABSTRACTS. Noncritical abstracts of the world's literature in psychology and related subjects. Monthly. Approximately 20,000 abstracts published annually. The 1969 volume is Vol. 43. Subscription: \$40.00 (foreign \$40.50). Single copy \$4.00.

PSYCHOLOGICAL BULLETIN. Evaluative reviews of research literature, discussions of research methodology in psychology. Monthly. Two volumes per year, six issues in each volume. The 1969 volumes are Vols. 71 and 72. Subscription: \$20.00 for 2 vols. (foreign \$20.50). Single copy \$2.00.

PSYCHOLOGICAL REVIEW. Original contributions of a theoretical nature. Bimonthly; first issue appears in January. The 1969 is Vol. 76. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

NEW IN 1969 DEVELOPMENTAL PSYCHOLOGY. Studies on growth and development and their associated variables. Bimonthly; first issue appears in January. The 1969 volume is Vol. 1. Subscription \$10.00 (foreign \$10.50). Single copy \$2.00.

Discounts: AGENTS and BOOK DEALERS receive a 10% discount on all orders. There is no club arrangement. There are no quantity, cash, library, or other special discounts. *Payment in U. S. funds.*

★ Subscriptions are available on a January through December basis only ★

AMERICAN PSYCHOLOGICAL ASSOCIATION
Subscription Department
1200 Seventeenth Street, N.W.
Washington D. C. 20036

JOURNAL OF CONSULTING and CLINICAL PSYCHOLOGY

August 1969

Vol. 33, No. 4

Contents

ARTICLES

- Follow-Up of 812 Girls 10 Years after High School Dropout: *Starke R. Hathaway, Phyllis C. Reynolds, and Elio D. Monachesi* 383
- Bodily Concerns and the WISC Object Assembly Subtest: *Gerald Marsden and Neil Kalter* 391
- Situational Appraisal Inventory: Development and Validation of a Measure of Evaluative Attitudes: *Stephen M. Pittel and Gerald A. Mendelsohn* 396
- Extreme Response Tendency as a Function of Emotional Adjustment and Stimulus Ambiguity: *Russell P. Norman* 406
- Personality Factors Related to Response to Social Isolation and Confinement: *Dalmas A. Taylor, Irwin Altman, Ladd Wheeler, and Estelle N. Kushner* 411
- Psychological Characteristics of Neighborhood Youth Corps Enrollees: *Sherman Eisenthal and Lewis J. Sherman* 420
- Effects of Pre-session Information on Perception of the Therapist and Receptivity to Influence in a Psychotherapy Analogue: *Roger P. Greenberg* 425
- Digit Span: An Indicant of Trait or State Anxiety?: *William F. Hodges and Charles D. Spielberger* 430
- Client-Therapist Relationship Factor in Systematic Desensitization: *Jon E. Krapfl and M. Mike Nawas* 435
- Effects of Vicarious Therapy Pretraining and Alternate Sessions on Outcome in Group Psychotherapy with Outpatients: *Charles B. Truax and Donald G. Wargo* 440
- Measurement of Social-Evaluative Anxiety: *David Watson and Ronald Friend* 448
- Effect of Physical and Cognitive Somatic Arousal on Rorschach Responses: An Experimental Test of the Assumption that Body Image Influences the Perceptual Organization of Unstructured Stimuli: *Steven N. Van De Mark and Charles Neuringer* 458
- Suicide and Natural Death in a State Hospital Population: A Comparison of Admission Complaints, MMPI Profiles, and Social Competence Factors: *Milton R. Ravensborg and Adeline Foss* 466
- Females' Affective Responses to Reading Erotic Literature: *Donald L. Mosher and Irene Greenberg* 472
- Object-Discrimination Learning Set Acquisition in Young Cerebral Palsied Children in Relation to Tested and Rated Intelligence: *Douglas Pearson* 478
- Internal-External Control Dimension as Related to Accident and Suicide Proneness: *Christene Blanton Williams and James B. Nickels* 485
- Child-Parent Identification and Academic Underachievement: *Elnore Granlund and Lois Knowles* 495
- Word Association Repertoires of Schizophrenics and Normals: *George D. Fuller and Solis L. Kates* 497
- Intelligence of Volunteers as Research Subjects: *Jay M. Cudrin* 501
- Humor and Suicide: Favorite Jokes of Suicidal Patients: *Don Spiegel, Patricia Keith-Spiegel, Joel Abrahams, and Lionel Kranitz* 504

BRIEF REPORTS

- Motivational Properties of Personality Conflict: *Ronald E. Smith* 506
- Perceived Maternal Nurturance and Control of Process Schizophrenics, Reactive Schizophrenics, and Normals: *Peter A. Magaro and Barbara A. Hanson* 507
- Performance Characteristics of School Children on the BIP Bender Test: *Jerry Adams and Arthur Canter* 508

LIST OF MANUSCRIPTS ACCEPTED 494, 503

Brief Reports

The *Journal of Consulting and Clinical Psychology* will accept Brief Reports of research studies in clinical psychology for early publication without expense to the author. The procedure is intended to permit the publication of soundly designed studies of specialized interest or limited importance which cannot now be accepted because of lack of space. Several pages in each issue will be devoted to Brief Reports, published in the order of their receipt without respect to the dates of receipt of the regular articles. Most Brief Reports appear in the first or second issue to go to press following their final acceptance.

An author who wishes to submit a Brief Report:

1. Sends the Brief Report, limited to one printed page and prepared according to the specifications given below.
2. Also sends to the Editor a full report of the research study, in sufficient detail to give a clear account of its background, procedure, results, and conclusions, which will be filed with the National Auxiliary Publications Service (NAPS) to insure indefinite availability.
3. Prepares at least 100 mimeographed copies of the full report, which the author will send without charge to all who request it as long as the supply lasts.
4. Agrees not to submit the full report to another journal of general circulation.

Specifications

Brief Report. The Brief Report should give a clear, condensed summary of the procedure of the study and as full an account of the results as space permits.

To insure that the Brief Report will be no longer than one printed page, its typescript, in-

cluding all matter except the title and the author's lines, must not exceed 85 lines averaging 42 characters and spaces in length. Set the type-writer margins for short lines of 42 characters, which are 3.5 inches long in elite typing, and 4.2 inches long in pica.

The manuscript of the Brief Report must be *double spaced* throughout. Except for its short lines, it follows the standard style of the 1967 revision of the *APA Publication Manual*. Headings, tables, and references are avoided or, if essential, must be counted in the 85 lines. Each Brief Report must be accompanied by a footnote in the style below, which is typed on a separate sheet and *not* counted in the 85-line quota:

¹ An extended report of this study may be obtained without charge from John Doe (giving the author's full name and address) or for a fee from the American Society for Information Science. Order NAPS Document — from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 909 Third Avenue, New York, New York 10022; remitting \$ — for microfiche or \$ — for photocopies.

Extended report. Because the extended report is intended for photoduplication, and is not copy to be sent to a printer, its style should differ in several ways from that of other manuscripts: (a) The extended report should be typed with single spacing for economy in duplication. (b) Tables and figures should be placed adjacent to the text which refers to them. A caption should be typed below each figure. (c) Footnotes should be typed at the bottom of the page on which reference is made to them. In other respects, the full report is prepared in the style specified by the *Publication Manual*.

Style of Manuscripts. Manuscripts must be prepared in the style described in the *Publication Manual of the American Psychological Association (1967 Revision)*, obtainable for \$1.50 from the Washington office of the APA.

Early Publication. Authors may secure the early publication of an acceptable article in the first issue to go to press after its acceptance by arranging for the article to be printed as extra pages, increasing the number of pages that subscribers receive. Information about charges for early publication may be obtained from the Managing Editor.

National Auxiliary Publications Service. To save printing costs, materials supplementary to articles may be deposited with the NAPS. To obtain copies, order by Document Number, as cited in the reference footnote, from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001. The indicated fee must be remitted in advance.

Abstracts. Manuscripts must be accompanied by an abstract of 100–120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or they may be obtained from the Editor or from the APA Central Office.

Copyrighted material. The following policies govern reprinting or other forms of reproduction of APA-copyrighted journal materials. Permission by the American Psychological Association must be requested for reprinting all tables, figures, and text of more than 500 words (total from one article) in length. There will be a charge of \$5.00 per page for such permission except when the requester is the author of the article to be reprinted, or the article is to be reproduced in a limited number of copies solely for instructional purposes.

Reprints. Authors receive 50 reprints gratis. Additional quantities may be ordered when returning proofs.

FOLLOW-UP OF 812 GIRLS 10 YEARS AFTER HIGH SCHOOL DROPOUT¹

STARKE R. HATHAWAY,² PHYLLIS C. REYNOLDS, AND ELIO D. MONACHESI

University of Minnesota

A dropout group and a nondropout group were compared on several variables. The data both predated and postdated the time of school dropout. For analysis the groups were subdivided on time of marriage and intelligence. Dropouts have a higher separation and divorce rate, even when time of marriage is controlled; they have larger families, but apparently only because they have been married longer. In general they marry into a lower socioeconomic status, but only girls from blue collar families show a downward social mobility. When grouped by intelligence, high and medium intelligence dropouts do not show the upward social mobility that nondropouts do. Dropout for girls seems less catastrophic than it does for boys.

Rich man, poor man, beggar man, thief, doctor, lawyer, merchant, chief—little girls count on their buttons to predict whom they will marry when they grow up, the prediction of the day depending on the number of buttons on their clothes. Even if a little girl were enterprising enough to search the psychological and sociological literature, she could not really improve on her button method; there is a paucity of studies on the social mobility of women. It is known, for example, that dropping out of high school has a real effect on the socioeconomic status of boys; the social mobility is in the downward direction for these boys by 10 years after quitting high school (Hathaway, Reynolds, & Monachesi, 1969). It has not been known, however, if there is a similar effect on the socioeconomic status of dropout girls by the time they marry, although Wolfbein (1959) shows that more girl dropouts have unskilled first jobs or are unemployed compared to high school graduates 1-5 years after quitting school.

Much has been written recently about the dropout (Cervantes, 1965; Greene, 1966; National Committee for Children and Youth, 1961; Schreiber, 1964, 1967); however, in practically all of this literature, female drop-

outs are either excluded or are included but not distinguished from male dropouts. As Deutsch (1964) says,

it is often in school that another element of the dropout problem related to another type of discrimination takes place—and it is an element that is often ignored. That is, the female dropout. Too often, discussion of dropouts deal with all cases as an undifferentiated totality or concentrate on males, without recognizing that many of the elements responsible for such a high dropout rate among American Indians, Negro Americans, etc., are similar in the case of females [p. 213].

The present paper explores some variables that distinguish girls who drop out of high school from girls who graduate, and compares the fate of the female dropout to that of the male dropout.

We were unable to develop the findings from the Minnesota Multiphasic Personality Inventory (MMPI) and certain other variables that were analyzed in the report of the careers of dropout boys (Hathaway et al., 1969). We do, however, have these data on girls and expect to publish them in another report.

PROCEDURE

Samples

The dropout and control samples for this study were drawn from 11,329 Ss who comprised the statewide sample in a longitudinal study of adolescent personality and behavior which began in 1953 when the Ss were ninth-graders in Minnesota schools. The schools were selected in 1953 to represent Minnesota's diverse economic and geographic areas. The statewide sample was a stratified random one except

¹ This research was supported by Research Grant 5 R01 MH 12187 from the National Institute of Mental Health and grants from the University of Minnesota.

² Requests for reprints should be sent to Starke R. Hathaway, Medical School, Division of Clinical Psychology, University of Minnesota, Minneapolis, Minnesota 55455.

that it did not include either Minneapolis or St. Paul. It did include 13 Twin City suburbs of varying social status level and the city of Duluth. It was impossible to record race when the data were gathered; because of the racial composition of Minnesota, however, it is safe to say that the sample represents a mainly white, midwestern society. For a more detailed description of the data collection and other generalities on this longitudinal study see Hathaway and Monachesi (1963).

The original number of girls was 5,628, of whom 816 (14%) dropped out of school before graduation. Because four of the dropouts could not be found at follow-up time, the basic total dropout *N* for this paper was 812. The control sample was 812 girls who graduated from high school. These girls were selected on a random basis within each school to match in number the dropouts from that school.

Variables

During the high school years 1953-1957 data were gathered on the following variables pertinent to this paper: socioeconomic status (SES), as defined by the Minnesota Scale for Paternal Occupations (Institute of Child Welfare, undated); intelligence test results based on ninth- or twelfth-grade ACE or, in a few instances, IQ equivalents; reason for drop out; and time of drop out.

During a 1962-1966 follow-up, data were gathered on year of marriage, present marital status, number of children, occupation and SES of husband, and number of school years eventually completed.

Although 94% of the original sample was contacted in some way, because of the usual difficulties inherent in a longitudinal study the numbers of lost cases on follow-up variables ranged from 29% (SES of husband) to 11% (present marital status) for the dropouts; the similar attrition of numbers for the controls was proportionately less because, as a group, the controls were easier to locate at follow-up time.

RESULTS

Data published earlier (Hathaway & Monachesi, 1963) showed that girls who drop out of school significantly more frequently come from blue collar and less frequently from white collar families than do girls who complete high school, that they more frequently come from broken families, and that they more frequently have low intelligence test scores. Among the 82% of dropout girls for whom a reason for drop out was known, 38% gave marriage as the reason. This was the most frequent single reason. Twenty-two percent professed lack of interest, and 15% were reported to be pregnant out of wedlock (Pauker, 1968). Nine percent dropped out to work. Only 5% dropped out with failure in school as the apparent reason.

At the end of the follow-up, only 5% of the dropouts had remained single, which contrasted to 16% of the controls who were still single. Comparison of the married groups showed that the rate of separation or divorce was significantly higher for dropouts than for controls. When dropouts alone were grouped by time of marriage (before June 1957 and after June 1957), there was no significant difference in divorce rate between the two groups. But when those dropouts married after June 1957 were compared to controls married after June 1957, the divorce rate for the dropouts was significantly higher.

The dropouts who married early had significantly larger families than later dropouts: 37% of the early marrieds had four or more children compared to 9% of the after-June 1957 group. When all dropouts were compared to controls, there was also a significant difference in family size: 27% of all dropouts had four or more children compared to only 8% of all controls; only 6% of the dropouts had no children, compared to 19% of the controls. When only those dropouts and controls who married after June 1957 were compared, however, there was no significant difference in family size. Dropping out of school does not seem to be an indication of high fecundity; as a whole, dropouts marry earlier than graduates and therefore have a greater opportunity to produce children. Also, as indicated above, pregnancy is a determiner of dropout so the early marriage group, at least, is biased.

The Hathaway and Monachesi (1963) study did not show evidence of any clear-cut linear relationship between time of dropout and severity of social problems or intelligence. Similarly, the data of the present study did not show significant relationships between time of marriage (early vs. later) of dropouts and the variables of socioeconomic level of the parents, present socioeconomic status, parents' marital status, or intelligence test scores. As stated above, number of children produced the only significant difference. Similarly, and in line with the earlier results, no significant relationship was found between time of dropout and present socioeconomic level.

Because intelligence test scores were found to differentiate dropouts from controls in the earlier study, the dropouts were grouped by high (69-100 percentile), medium (40-68 percentile), and low (1-39 percentile) aptitude scores and compared on several of the variables. Among these dropout groups, the intelligence levels showed no significant relationship to years of school eventually completed (less than high school vs. eventual high school graduation or higher), early or later marriage, present marital status, present socioeconomic level, or number of children.

When both the dropouts and the controls were divided by intelligence and high was compared to high, medium to medium, and low to low, a different pattern appeared. Both high and medium intelligence groups showed significant frequency differences in marital status categories, but in both cases the largest contribution to the significance level came from the different percentage of single girls in each group. As was reported above, the controls included a much higher frequency in this category. When the unmarried women were eliminated from the samples and the still-marrieds were compared to the separated or divorced girls, the high intelligence dropouts had a greater separation and divorce rate than high intelligence controls; between the medium intelligence groups this difference was less reliable, and differences between low intelligence dropouts and low intelligence controls fell to an unreliable level. All three aptitude level dropout groups had significantly larger families when compared to their control groups. This may relate to the selection factor mentioned above for family size.

Dropouts differed significantly from controls on present socioeconomic level (based on husband's occupational level). Table 1 gives the percentages in each SES category. The differences in the slightly skilled and the unemployed categories are not meaningful. Even with these relatively large N s, the SES levels had to be grouped into: (a) white collar (professional, semiprofessional, clerical, skilled); (b) blue collar (semiskilled, slightly skilled); (c) day laborer and unemployed; and (d) farmer, for other comparisons. With these

TABLE 1

OCCUPATIONAL LEVEL OF HUSBANDS OF DROPOUTS
COMPARED TO THOSE OF CONTROLS

Occupational level	Dropouts	Controls
Professional and semi-professional	4	21
Clerical, skilled	21	23
Semiskilled	30	26
Slightly skilled	24	17
Day laborer	5	1
In service (only)	6	5
Farmer	7	5
Unemployed	3	11
Total	100	99

Note.—All entries are in percentages. For dropouts, $N = 481$; for controls, $N = 491$. $\chi^2 = 82.61$, $df = 7$, $p < .001$.

four classes, significant differences in SES level continued to appear when the dropout and control samples were grouped by intelligence. The high intelligence and medium intelligence controls showed significant differences. For the high intelligence groups the difference was most apparent at the white collar level: far fewer than the expected number of dropouts married men who had white collar jobs. For the medium intelligence groups the difference was more at the day laborer and unemployed level, with twice the expected number of dropout girls who married men at this lower occupational level. The comparison of low intelligence dropouts to low intelligence controls produced no statistically significant difference. These findings are shown in Table 2 which gives the percentages in each SES category for each intelligence group. Whereas the difference between the high and low intelligence control group in the white collar category was 26%, the difference in the same category for the dropout group was only 8%. Similarly, the control group blue collar category difference from high to low intelligence was 21%, and for the comparable dropout group was only 5%. The farmer category was almost constant across the table.

Although the data showed that girls who drop out of high school tended as a group to come from lower SES levels and to continue by marriage at lower SES levels, a study of social mobility following each girl could show if, among dropouts, there was really upward

TABLE 2

PERCENTAGES OF MARRIED DROPOUTS AND CONTROLS IN EACH SES CATEGORY (BASED ON HUSBAND'S OCCUPATION) WHEN GROUPED BY INTELLIGENCE PERCENTILES

SES category	High intelligence ^a		Medium intelligence ^b		Low intelligence ^c	
	Dropout (N=103)	Control (N=180)	Dropout (N=128)	Control (N=128)	Dropout (N=203)	Control (N=156)
White collar	31	58	28	48	23	32
Blue collar	56	36	52	44	61	57
Day laborer or unemployed	6	1	12	0	9	5
Farmer	7	5	9	8	7	5
Total	100	100	101	100	100	99

^a $\chi^2 = 20.22$, $df = 3$, $p < .001$.

^b $\chi^2 = 23.95$, $df = 3$, $p < .001$.

^c $\chi^2 = 5.10$, $df = 3$, $p < .25$.

or downward mobility. The SES level of the family of origin of each married dropout and control *S* was compared to the occupational level of *S*'s husband. When compared to the controls by this procedure, only the dropouts whose parents were at the blue collar level showed a significant downward mobility trend. Far fewer than the expected number of girls married into the white collar class and far more than the expected number married men who were day laborers or unemployed. The percentages of girls moving from blue collar families into each SES category are given in Table 3.

None of the other SES mobility data yielded significant differences. Thus, if a girl who drops out of school comes from a white collar family, she is just as likely to marry a man with a white collar job as a girl from the same SES level who finishes school.

Another look at SES mobility was provided from grouping by intelligence test percentiles. For these analyses coarser SES grouping had to be used because of small *N*s. The blue collar, day laborer, and unemployed categories were combined. As with other analyses where the girls were grouped by intelligence, it was the high and medium intelligence dropouts who contributed the chief differences. The low intelligence dropouts did not differ significantly from the low intelligence controls in any of the SES movement categories. The high intelligence dropouts from blue collar or below SES families did not show a reliable trend to upward mobility as did their comparable

control groups (χ^2 , $p < .001$). This same relationship was found in the mobility trends between medium intelligence dropouts and medium intelligence controls, but with lesser certainty (χ^2 , $p < .01$). This same relationship was found in the mobility trends between medium intelligence dropouts and medium intelligence controls, but with lesser certainty (χ^2 , $p < .01$). Because of the necessary coarse grouping of blue collar with day laborer and unemployed, a downward trend, if it existed, could not be demonstrated. Table 4, which compares the SES mobility between low and high intelligence groups for both the dropouts and controls, shows that, for the controls, intelligence seems to have some effect on SES level; a high intelligence girl from a white collar family tends to marry a man at the white collar level, or if she comes from a blue

TABLE 3

MOVEMENT OF DROPOUTS AND CONTROLS FROM BLUE COLLAR FAMILY OF ORIGIN TO FAMILY OF PROCREATION LEVEL

Item	Dropouts (N=193)	Controls (N=207)
Parent to husband		
Blue collar to white collar	26	46
Blue collar to day laborer/ unemployed	60	47
Blue collar to farmer	6	5
Total	101	99

Note.—All entries are in percentages. $\chi^2 = 27.19$, $df = 3$, $p < .001$.

TABLE 4

COMPARISON OF SES MOBILITY BETWEEN LOW AND HIGH INTELLIGENCE DROPOUTS
AND LOW AND HIGH INTELLIGENCE CONTROLS

SES mobility	Dropouts		Controls	
	Low intelligence	High intelligence	Low intelligence	High intelligence
Parent to husband				
White collar to white collar	39	53	40	67
White collar to blue collar/day laborer/unemployed	61	47	60	33
Total	100	100	100	100
<i>N</i>	31	19	25	61
Blue collar/day laborer/unemployed to white collar	23	22	34	56
Blue collar/day laborer to unemployed	77	78	66	44
Blue collar/day laborer/unemployed				
Total	100	100	100	100
<i>N</i>	105	54	91	80

Note.—All entries except *N* are in percentages.^a $\chi^2 = 1.37$, $df = 1$, $p < .25$.^b $\chi^2 = 5.86$, $df = 1$, $p < .025$.^c $\chi^2 = 0$.^d $\chi^2 = 7.70$, $df = 1$, $p < .01$.

collar level, she tends to marry above the level of her family of origin. These tendencies are not apparent for the girl who drops out of high school; a high intelligence girl from a white collar family is just as likely as a low intelligence girl from a white collar family to marry in the same broad social class.

As a final analysis of social mobility, data were taken from the authors' report on dropout boys (Hathaway et al., 1969) for comparison to similar data on the girls. Table 5 presents these comparisons for male and female dropout groups and the control groups. Although differences are not marked, the pattern for male and female controls seems to be that of stable to upward social mobility, but when male and female dropouts are compared, a downward trend for the boys is apparent.

Table 6 gives the actual outcome proportions of male and female dropouts in each SES category, that is, the level of the *S*, if a male, or the spouse, if a female. Clear differences between males and females appeared at both extremes—more than the expected number of dropout females married into the professional/semiprofessional level and more than the expected number of dropout males became day laborers.

DISCUSSION

Girls in Minnesota who drop out of high school before graduating can be characterized as of average to below-average intelligence (only 22% have intelligence test scores above the 68th percentile) and as coming from low-income families (66% of the parents are at the semiskilled level or below). They marry readily and early, and, when compared to girls who graduate from high school, they tend to marry men with low income jobs (62% of the husbands are at the semiskilled level or below). Among these dropout girls, however, being bright or not so bright seems to have little relationship to their eventual socioeconomic level. Ten years later, the dropouts average out with larger families than nondropouts mainly because many of them have been married longer, but their separation and divorce rate is higher even when compared to nondropouts who have been married the same length of time.

If socioeconomic level of the family she establishes can be used as a criterion of success in life, social mobility data can help to answer the question of how much effect school dropout has on a girl's later chances of success. A comparison with nondropouts of

the same parental SES level shows that dropout girls who come from blue collar (semi-skilled and slightly skilled) families, and only this group, have a tendency to downward mobility; 60%, however, remain at the same blue collar level when they marry. When compared to their control group, dropout girls from white collar families show no downward trend, and dropout girls from day laborer and unemployed families show an upward mobility as much as does their control group.

TABLE 5

SES MOBILITY COMPARISONS BETWEEN MALE AND FEMALE DROPOUT GROUPS AND THE CONTROL GROUPS

SES mobility	Dropouts		Controls	
	Female	Male	Female	Male
Parent to husband				
White collar to white collar	40	23	57	61
White collar to blue collar	54	63	42	34
White collar to day laborer/unemployed	6	14	1	5
Total	100	100	100	100
N	70	103	112	168
Blue collar to white collar	27	20	49	44
Blue collar to blue collar	63	63	50	48
Blue collar to day laborer/unemployed	9	17	1	8
Total	100	100	99	100
N	182	262	196	303
Day laborer/unemployed to white collar	24	18	41	35
Day laborer/unemployed to blue collar	64	60	51	54
Day laborer/unemployed to day laborer/unemployed	12	21	8	10
Total	100	99	100	99

Note.—All entries except N are in percentages.

^a $\chi^2 = 6.65$, $df = 2$, $p < .05$.

^b $\chi^2 = 2.59$, $df = 2$, $p < .50$.

^c $\chi^2 = 7.92$, $df = 2$, $p < .025$.

^d $\chi^2 = 11.27$, $df = 2$, $p < .005$.

^e $\chi^2 = .13$, $df = 2$, $p < .50$.

^f $\chi^2 = 3.09$, $df = 2$, $p < .25$.

TABLE 6

COMPARISON OF OUTCOME SES LEVEL OF FEMALE AND MALE DROPOUTS

Outcome SES level	Female	Male
Professional and semiprofessional	4	2
Clerical, skilled	22	17
Semiskilled	31	27
Slightly skilled	26	31
Day laborer	6	11
Farmer	8	7
Unemployed	3	5
Total	100	100

Note.—All entries are in percentages. For females, $N = 453$; for males, $N = 680$. $\chi^2 = 23.93$, $df = 6$, $p < .001$.

In general, high school dropout for girls does not have the severe ramifications that it has for boys. Similar to the girls, boy dropouts tend to come from low-income families, but unlike the girls, their social mobility direction is downward regardless of the parental SES level. It is only at the day laborer, unemployed level where boy dropouts do not differ significantly from girl dropouts in mobility direction, although the downward trend for boys is still apparent. Among male dropouts, 57% have intelligence test scores at or below the 39th percentile; among the female dropouts, the percentage is 47. When each of these groups is compared to its control group, that is, boys or girls of the same intelligence level who graduated from high school, the data show that dropout girls as a group do not suffer the consequences that dropout boys do regarding eventual socioeconomic level (81% of the dropout boys had jobs at the semiskilled level or below, compared to 56% of the control boys; 70% of female low-intelligence dropouts married at the semiskilled level or below, compared to 62% of the control low-intelligence group). These findings can be interpreted as an indirect indication of the importance of the high school diploma for boys and the lesser importance it has for girls—the most “justifiable” dropouts would be the ones who lack the intelligence to finish the academic curriculum of the average high school, and girls whose dropping out is thus “justified” fare no worse than girls of com-

parable intelligence who persevere to the end of high school.

Another group of girls who should be considered, however, are those of high intelligence. Do these girls, who have the mental capacity to continue on to the high school diploma and beyond, lose out in the long run? And does society lose out because of wasted talent? Twenty-two percent of the girl dropouts were of high intelligence. When these girls are compared to girls of high intelligence who finish high school, a clear case of handicap is demonstrated: only 5% of the dropouts eventually went back to school to get their high school diploma, and of these, one girl had 1 year of college, one girl finished college, and one had 1 year of postgraduate work. Gibson, Higgins, and Mitchell (1967) found that 11% of high school dropouts (both male and female) in Toledo had "potential for college." After contact and invitation to participate in a workshop that led to the university, only 2% of this group completed their freshman year. These authors conclude that "more fruitful efforts might be directed toward 'talent retention' rather than 'talent reclamation.'" The present data suggest that even without a reclamation process, 2% of dropout girls with potential go on to college, casting doubt on the value of a reclamation program, yet concurring with the Gibson et al. conclusion. As a reminder of contrasting careers, 65% of the high school graduates went on to college or training (e.g., nursing), and of these, 35% finished college.

Among the high intelligence dropouts, 97% had been married by the time of follow-up, and of these, 13% were either separated or divorced. Again, for contrast, 81% of the high school graduates were married, and of these, 5% were separated or divorced. Twenty-seven percent of the high intelligence dropouts had four or more children, but only 5% of the graduates had as many. Only 31% of the dropouts married at the white collar level (only 7% at the professional, semiprofessional level) compared to 58% of the high school graduates (31% at the professional, semiprofessional level). Only 20% of the dropout girls came from white collar families (compared to 35% of the graduates); the dropouts, how-

ever, did not marry below themselves in social class any more than did the graduates. It appeared to be the dropouts from the blue collar and lower SES levels whose social mobility was downward in comparison to the graduates.

If talent existed among these high intelligence dropouts, it now seems to lie dormant in typically, a young lady who is married, whose husband has a semiskilled job, who has several children and who, perhaps because of lack of time, has never obtained her high school diploma.

CONCLUSION

Few studies have been reported on girls who drop out of high school before graduation. Perhaps researchers have suspected what is in part borne out by this paper: the consequences of drop out for girls are not so severe as they are for boys. Future research on dropouts should study each sex separately. In the 1950s, only 14% of girls in Minnesota quit school before graduation. Ten years later practically all of these girls had married and many had produced sizable families. Their social mobility was more stable than for boys who dropped out of school. Both boy and girl dropouts tended to, in the authors' data, come from and continue at the lower socioeconomic levels. Very few of the girls returned to school to obtain a diploma. As a group they had a higher separation and divorce rate than the girls who graduated. This was true even when a correction was made to equate the time of marriage.

REFERENCES

- CERVANTES, L. F. *The dropout: Causes and cures*. Ann Arbor: University of Michigan Press, 1965.
- DEUTSCH, M. Early social environment: Its influence on school adaptation. In D. Schriber (Ed.), *The school dropout*. Washington, D. C.: National Education Association, 1964.
- GIBSON, R. L., HIGGINS, R., & MITCHELL, M. H. The high school dropout goes to college. *Personnel and Guidance Journal*, 1967, 45, 824-827.
- GREENE, B. I. *Preventing student dropouts*. Englewood Cliffs, N. J.: Prentice-Hall, 1966.
- HATHAWAY, S. R., & MONACHESI, E. D. *Adolescent personality and behavior*. Minneapolis: University of Minnesota Press, 1963.

- HATHAWAY, S. R., REYNOLDS, P. C., & MONACHESI, E. D. Follow-up of the later careers and lives of 1,000 boys who dropped out of high school. *Journal of Consulting and Clinical Psychology*, 1969, 33, 370-380.
- INSTITUTE OF CHILD WELFARE, UNIVERSITY OF MINNESOTA. *Minnesota Scale for Paternal Occupations*. Minneapolis: Author, undated.
- NATIONAL COMMITTEE FOR CHILDREN AND YOUTH. *Social dynamite: The report of the conference on unemployed, out-of-school youth in urban areas*. Washington, D. C.: Government Printing Office, 1961.
- PAUKER, J. D. Girls pregnant out of wedlock: Are they pregnant because they are different, or are they different because they are pregnant? Paper presented at the National Conference on Social Welfare, San Francisco, May 1968.
- SCHREIBER, D. (Ed.) *The school dropout*. Washington, D. C.: National Education Association, 1964.
- SCHREIBER, D. (Ed.) *Profile of the school dropout*. New York: Random House, 1967.
- WOLFBEIN, S. L. Transition from school to work: A study of the school leaver. *Personnel and Guidance Journal*, 1959, 38, 98-105.

(Received July 26, 1968)

BODILY CONCERNS AND THE WISC OBJECT ASSEMBLY SUBTEST

GERALD MARSDEN¹ AND NEIL KALTER²

University of Michigan

Previous examinations of the hypothesis that bodily concerns are reflected in impaired WISC Object Assembly performance have produced contradictory results, but have also used different criteria for assignment of Ss to two *S* groups. To investigate whether these differences account for the discrepant results, 73 white, latency and early adolescent boys were assigned to three groups differing in intensity of bodily concerns. One-way analyses of variance with covariance adjustments for relevant variables failed to yield any support for the hypothesis. Theoretical limitations of the hypothesis were discussed and a more appropriate approach suggested and supported by *prima facie* evidence gleaned from analysis of the finding that intensely bodily concerned boys performed significantly better ($p < .01$) than others on the Information subtest.

Blatt, Allison, and Baker (1965) found support for the hypothesis that concerns about body intactness and integrity are reflected in the Wechsler intelligence tests by a subtest pattern in which the Object Assembly (OA) subtest score is lower than the scores of many of the other subtests. Among the varieties of data brought to bear on this hypothesis was a comparison of two groups of children who differed in concerns about body intactness. Records of children previously evaluated at a child guidance center were reviewed by a clinical psychologist, not familiar with the hypothesis, who assigned them either to a group clearly expressing bodily concerns or to a group for which bodily concerns were not believed to be a dominant issue. The Wechsler Intelligence Scale for Children (WISC) OA subtest scaled scores of the children in these groups were then compared, and found to differ, as expected, to a degree falling just short of the .05 level of significance. When the groups were compared in terms of the deviation of OA scaled scores from mean scaled scores computed across all 11 subtests and across the five Performance Scale subtests, the expected difference was

again found, this time significant in both analyses at the .03 level. Similar analyses of the other WISC subtests yielded no significant differences.

Rockwell (1967) criticized this portion of the Blatt et al. study on the grounds that the *N* was too small to permit generalization (the two groups totaled 13 Ss), that Ss in these groups were apparently not matched in any way, that the groups were biased with respect to sex (three girls were included in the two groups), and that the bodily concern group contained one *S* with a very low Full Scale IQ (FSIQ) score (when this *S* was removed in Rockwell's reanalysis of the Blatt et al. data the significant relation reported above failed to appear). Rockwell then replicated the study. He selected two groups of 15 boys each from the files of the University of Minnesota Child Development Clinic. They were matched within two IQ points on WISC FSIQs and all had IQ scores of 90 or above. Rockwell found no significant relation between the bodily concern groups on OA or on any other subtest when he repeated the analyses employed by Blatt et al.

One cannot be certain, however, that Rockwell's effort constituted a replication of the Blatt et al. study. The point at issue has to do with the criteria for assigning Ss to the two groups. Blatt et al. appear to have included in the bodily concern group children for whom bodily concerns were a *dominant* issue, and in the control group children who,

¹Requests for reprints should be sent to Gerald Marsden, Children's Psychiatric Hospital, University of Michigan Medical Center, Ann Arbor, Michigan 48104.

²The authors gratefully acknowledge the statistical consultation provided by J. E. Keith Smith, and the helpful criticism of an earlier draft by Albert C. Cain, Martin Mayman, and Paul Meisel.

while they may have had some degree of concern about bodily intactness, were more concerned about other issues. Rockwell, on the other hand, included in his bodily concern group children whose records included the statement that bodily concerns were present, while the records of the children in his control group contained no mention of bodily concerns. Thus Rockwell's bodily concern group may well have contained children who would have been assigned to either of the Blatt et al. groups. And the Blatt et al. control group may have contained some children that Rockwell would have assigned to his bodily concern group. The question to be examined in this report is whether these apparent differences in the composition of the groups can account for the discrepancies in the findings reported by Blatt et al. and Rockwell.

METHOD

Records of children who had been evaluated at Children's Psychiatric Hospital³ were checked by a secretary to screen out those in which a complete WISC record was unavailable. The remaining records, including social work, psychiatric, and psychological reports (but not psychological test protocols), and treatment notes were then drawn alphabetically for study by the junior author. He was, at that time, not familiar with the hypothesis, having been told only that the study involved a possible relation between certain WISC subtests and bodily concerns. His task was to eliminate all cases in which there was question of organicity or in which the WISC had been administered previously, and to assign the remaining cases to one of three groups.

Group 1 was comprised of cases in which concern about bodily damage, integrity, or intactness represented the primary mode of giving expression to the concerns, conflicts, or impulses which lay at the heart of the psychopathology and presenting problems for which the child was being evaluated. In each of these cases two or more of the evaluators, particularly psychologists and psychiatrists, agreed that bodily concerns were of fundamental importance in explaining the psychopathology. Moreover, in every case the child and/or his parents gave explicit verbal expression to their awareness of these concerns and their intensity. Thus, two important features of the Group 1 cases were that the level of inference involved in classifying them was minimal, and that conscious recognition of the child's bodily concerns by himself or by members of his family were confirmed by diagnostic evaluations which weighed unconscious factors heavily. Group 2 consisted of cases in which bodily concerns were present, but did not

constitute the major mode of impulse or conflict expression. Other modes, such as phobias, compulsions, or intense anxiety, all relatively unrelated (at least overtly) to the body concerns, provided parallel and important means of expression. Cases assigned to this group varied from some in which body concerns were clearly present and important to those in which such concerns were mentioned only in passing. Records of Ss assigned to Group 3 contained no hint that the child entertained bodily concerns.

A total of 137 case records were read. Of these, 33 were discarded because the WISC had been administered previously or because the question of organicity was open. Three other records proved impossible to rate and were discarded. Records of female and Negro children occurred in our sample with the same relative infrequency as their occurrence in the hospital population. Because there were too few of either to consider the effects of sex and race separately, the sample was restricted to white males. This left a total of 73 cases distributed as follows: Group 1 = 10, Group 2 = 33, Group 3 = 30.

Early in the research 22 potential Ss were categorized⁴ and then later, after grouping guidelines had been elaborated and refined, they were independently recategorized by the junior author. A comparison of the grouping decisions on these Ss provides a rough—and probably minimal—estimate of the reliability of S grouping. Interjudge agreement was 82%. Three of the four disagreements occurred between Groups 1 and 2, and in each of these instances the junior author assigned the S to Group 2, thus indicating the stringency with which Group 1 Ss were selected.

RESULTS

A one-way analysis of variance revealed that the three groups did not differ significantly on any of the following variables: age (in months) when tested, FSIQ, Verbal Scale IQ, Performance Scale IQ, year tested. The latter was examined to check the possibility that the groups might reflect, in part, changes in hospital personnel's diagnostic orientation and sensitivity over time. The optional Digit Span subtest (Verbal Scale) had been administered to only about half of the Ss. Thus, the Verbal and Full Scale IQs of all Ss were not strictly comparable. In order to bring to light any group differences which might be masked by this factor, the contribution of Digit Span to IQ scores was prorated out and the groups were compared on the basis of these modified scores. Again, no significant differences were found. This modification is incorporated in all

³ University of Michigan Medical Center, Ann Arbor, Michigan.

⁴ Early work on grouping was done by Susan Golden.

TABLE 1

MEANS, STANDARD DEVIATIONS, AND *F* RATIOS OF SUBJECT GROUPS ON BACKGROUND VARIABLES

Variables	Group 1		Group 2		Group 3		<i>F</i> ^b
	\bar{X}	<i>SD</i>	\bar{X}	<i>SD</i>	\bar{X}	<i>SD</i>	
Full Scale IQ	103.80	14.77	102.82	14.92	98.93	12.00	.81
Full Scale IQ ^a	104.40	14.36	103.82	14.24	99.17	12.50	1.11
Verbal Scale IQ	104.50	15.15	100.76	14.31	96.27	13.49	1.55
Verbal Scale IQ ^a	105.70	14.25	101.58	13.95	96.67	14.39	1.84
Performance Scale IQ	102.10	15.67	104.52	14.23	102.13	11.93	.28
Age (months)	128.90	23.34	122.55	17.99	124.87	22.81	.37
Year tested	63.60	2.12	62.82	2.19	62.67	2.43	.64

Note.—For Group 1, *N* = 10; for Group 2, *N* = 33; for Group 3, *N* = 30.^a Modified by omission of Digit Span subtest in approximately 50% of the cases.^b *df* = 2/72. None of these ratios reach significance.

subsequent analyses. Table 1 presents the means, standard deviations, and *F* ratios for the three groups on each of these variables.

The absence of significant differences on these background variables cleared the way for examination of the bodily concerns-Object Assembly hypothesis. There were three questions: Did Group 1 have a significantly lower mean OA scaled score than Group 2 or Group 3? Was Group 1's mean OA scaled score among the lowest of their mean scaled scores? And were there significant differences between the groups on any other subtests? These questions were examined by means of a one-way analysis of variance with covariance adjustment for FSIQ. Age was not used as a covariate because an age control is built into the method of arriving at scaled scores and the IQ score. The correlation between age (in months) and FSIQ was .06. Group differences were sought on each of the 10 WISC subtests. In order to avoid equating the groups on a variable (i.e., FSIQ) which was in part composed of the variable on which group differences were sought (subtests), the FSIQ scores were first modified by removing the contribution of the subtest under examination in each instance. This was done by the usual method of prorating WISC scaled scores in all analyses employing FSIQ as a covariate.

Table 2 presents subtests means and standard deviations for each group, and the *F* ratios for differences between groups. These data appear first with no covariance adjust-

ment and then with the adjustment for FSIQ. There was no statistically meaningful difference between groups on any of the subtests except Information. Not only did Group 1 fail to differ from either of the other groups on the OA scaled score, but neither Group 1 nor Group 2 scored lower on OA than on most other subtests as Blatt et al. predicted. Indeed, the mean OA scaled score fell *above* the median in the distribution of group mean subtest scaled scores for all three groups.

Of course, scaled scores might not be the most appropriate data in terms of which to test the hypothesis. Certain items in the OA subtest might arouse sufficient anxiety in children with intense bodily concerns to interfere with their performance on that item. Such interference could be masked when item raw scores are summed and transformed into scaled scores. This possibility was explored through an analysis of variance of the OA item raw scores with covariance adjustments for age and FSIQ. While this analysis represented a post hoc examination of an insignificant finding and is, therefore, of dubious respectability, it was undertaken here for heuristic reasons with full recognition that any significant findings might well be chance effects. The analysis revealed that only Item 2 (horse) differentiated the groups at the .05 level (adjusted means for Groups 1, 2, and 3 were, respectively, 5.58, 5.15, and 5.17; *F* = 3.23, *df* = 2/68). Thus, the high body-concern group (Group 1) had a significantly *higher* mean item score than the other two

TABLE 2

WISC SUBTEST MEANS AND STANDARD DEVIATIONS WITH AND WITHOUT ADJUSTMENT FOR FSIQ,
AND *F* RATIOS, FOR THREE BODILY CONCERN GROUPS

Subtest	Covariate	Group 1		Group 2		Group 3		<i>F</i> ^a
		\bar{X}	<i>SD</i>	\bar{X}	<i>SD</i>	\bar{X}	<i>SD</i>	
Information	None	12.00	2.72	9.61	2.58	9.00	2.92	4.32*
	FSIQ	11.78	2.72	9.38	2.58	9.32	2.92	6.10**
Comprehension	None	10.30	2.90	10.52	3.21	9.03	2.37	2.61
	FSIQ	9.99	2.90	10.42	3.21	9.24	2.37	1.76
Arithmetic	None	10.40	2.50	9.58	2.84	9.30	2.56	.61
	FSIQ	10.10	2.50	9.46	2.84	9.53	2.56	.28
Similarities	None	10.80	3.40	10.76	3.73	10.43	3.63	.07
	FSIQ	10.17	3.40	10.55	3.73	10.87	3.63	.24
Vocabulary	None	10.90	2.70	10.36	3.26	9.63	2.66	.83
	FSIQ	10.42	2.70	10.16	3.26	10.02	2.66	.12
Picture Completion	None	10.40	2.29	11.76	3.17	11.03	2.86	.95
	FSIQ	9.94	2.29	11.60	3.17	11.36	2.86	1.84
Picture Arrangement	None	9.90	2.88	10.12	3.17	9.40	2.30	.05
	FSIQ	9.63	2.88	9.99	3.17	10.14	2.30	.15
Block Design	None	10.70	3.29	10.88	2.64	10.97	2.37	.04
	FSIQ	10.38	3.29	10.72	2.64	11.25	2.37	.73
Object Assembly	None	10.50	2.54	10.97	2.82	10.07	2.94	.76
	FSIQ	10.20	2.54	10.84	2.82	10.31	2.94	.46
Coding	None	8.60	4.29	9.00	3.12	9.17	2.82	.11
	FSIQ	8.41	4.29	8.84	3.12	9.40	2.82	.46

Note.—For Group 1, *N* = 10; for Group 2, *N* = 33; for Group 3, *N* = 30.

^a *df* with no covariates = 2/72, with FSIQ covaried *df* = 2/71.

* *p* < .05.

** *p* < .01.

groups, and the hypothesis was not supported. Moreover, while the magnitude of this difference was statistically significant, it appears psychologically trivial.

DISCUSSION

Despite the use of more carefully differentiated *S* groups in this study, no evidence could be found supporting the hypothesis that bodily concerns are reflected in impaired performance on the WISC OA subtest. And this was true whether the analysis was conducted on subtest scaled scores or on the level of individual items.

Several factors might account for this negative finding. These include the accuracy and adequacy with which bodily concerns were reflected in the closed case files used to define the criterion groups, and the possibility that the relation between bodily concerns and OA performance may be obscured in latency and early adolescence by rapidly changing organization of cognitive capacities or by heightened

drive-defense constellations in the face of puberty. Most of the evidence Blatt et al. adduced in support of the hypothesis was based on work with adults. It is important to note that the OA subtest itself poses partially different psychological problems for children and adults. While all four items of the Wechsler Adult Intelligence Scale OA subtest require the adult *S* to *discover* what object the cut up pieces form when assembled, the child *S* is *informed* what the object is before he begins work on the first two WISC OA items.

But far more significant are theoretical limitations of the hypothesis. One of Blatt et al.'s purposes in testing the bodily concerns hypothesis was to extend the understanding of the ego-psychological functions underlying the OA subtest. Working from Rapaport, Gill, and Schafer's (1968) rationale for the subtest as primarily involving visual-motor coordination, Blatt et al. asked what could be expected to disrupt this psychological func-

tion. The theoretical inadequacy of the hypothesis lies in this one-sided emphasis on disruption and the research strategy implied by that emphasis which led them to seek directly disrupting factors. A more adequate theoretical approach, one more fully consistent with psychoanalytic ego-psychology theory, would suggest that the presence of bodily concerns interacting with mediating variables could produce either enhanced or impaired performance on OA, or might not be limited in their effects to this subtest.

Examples of these mediating variables are patterns of defense or cognitive control principles, and stimulus thresholds required to evoke their use. There is now considerable experimental evidence of interindividual and intraindividual differences in the use of cognitive controls, and of relations between certain of these controls and both intellectual abilities and defensive styles (Gardner, Holzman, Klein, Linton, & Spence, 1959; Gardner, Jackson, & Messick, 1960; Gardner & Long, 1962; Guthrie, 1967; Holzman & Gardner, 1959). That individuals use quite different defense mechanisms in managing anxiety and other affects generated by roughly similar conflicts is a clinical commonplace. While many persons may deal with concerns about bodily integrity through primarily repressive and avoidant defenses, resulting in temporary disorganization of visual-motor functioning on the OA subtest, others might use isolation or counterphobic defenses, for example, producing temporarily heightened visual-motor performance.

A case in point may be our finding of a significant difference in performance on the Information subtest between Group 1 and Groups 2 and 3 (Table 2). That boys with intense bodily concerns should out-perform other boys on this subtest was unexpected and did not occur in either the Blatt et al. or Rockwell studies. It may be, therefore, a chance finding, but it may also reveal the effect of mediating variables. In an attempt to uncover the source of the Group 1 S's superior performance, one-way analyses of variance with covariance adjustment for age and FSIQ were computed for each subtest item. Only three items significantly differen-

tiated the groups: Number 16, "Who wrote 'Romeo and Juliet'?" ($F = 10.96, p < .01$); Number 20, "Where is Chile?" ($F = 22.65, p < .01$); Number 27, "What is a barometer?" ($F = 3.20, p < .05$) ($df = 2/68$). None of these is obviously related to the issue of bodily concerns. Notably, all three appear in the last half of the subtest and two occur in the last third. Group 1 Ss appear to differ from other Ss in that they succeed with certain more difficult, esoteric items. These Ss seem to have been either more attentive to information available in their environment or to have such information more available to memory or both. Those subtest items which do appear relevant to bodily concerns were among the least differentiating. If the relation between intense bodily concerns and high Information subtest scores is meaningful, it is not as a direct function of items in the subtest which pertain to the bodily concerns issue, but must reflect more indirect, more complex psychological processes.

REFERENCES

- BLATT, S. J., ALLISON, J., & BAKER, B. L. The Wechsler Object Assembly subtest and bodily concerns. *Journal of Consulting Psychology*, 1965, 29, 223-230.
- GARDNER, R. W., HOLZMAN, P. S., KLEIN, G. S., LINTON, H. B., & SPENCE, D. P. Cognitive controls: A study of individual consistencies in cognitive behavior. *Psychological Issues*, 1959, 1(Monog. 4).
- GARDNER, R. W., JACKSON, D. A., & MESSICK, S. J. Personality organization in cognitive controls and intellectual abilities. *Psychological Issues*, 1960, 2(Monog. 8).
- GARDNER, R. W., & LONG, R. I. Control, defence and centration effect: A study of scanning behavior. *British Journal of Psychology*, 1962, 53, 129-140.
- GUTHRIE, G. D. *Changes in cognitive functioning under stress: A study of plasticity in cognitive controls*. (Doctoral dissertation, Clark University) Ann Arbor, Mich.: University Microfilms, 1967. No. 67-13,541.
- HOLZMAN, P. S., & GARDNER, R. W. Leveling and repression. *Journal of Abnormal and Social Psychology*, 1959, 59, 151-155.
- RAPAPORT, D., GILL, M. M., & SCHAFER, R. *Diagnostic psychological testing*. Vol. 1. Chicago: Year Book Publishers, 1945. (Republished: rev. ed., New York: International Universities Press, 1968.)
- ROCKWELL, G. J. WISC Object Assembly and bodily concern. *Journal of Consulting Psychology*, 1967, 31, 221.

(Received July 2, 1968)

SITUATIONAL APPRAISAL INVENTORY:

DEVELOPMENT AND VALIDATION OF A MEASURE OF EVALUATIVE ATTITUDES

STEPHEN M. PITTEL¹

AND

GERALD A. MENDELSON

Mount Zion Hospital, San Francisco

University of California, Berkeley

The Situational Appraisal Inventory (SAI) was designed to meet the need for a measure of subjective moral (evaluative) attitudes. Previous instruments used for this purpose were found inadequate because of their psychometric properties or psychological assumptions. Two 100-item forms (SAI-A and SAI-B) were developed and standardized for college populations and a 30-item form (SAI-J) was derived from SAI-A to represent the cluster dimensions and total score of the longer form. Internal consistency and test-retest reliabilities of all forms exceeded .90 and all forms of the test were shown to be comparable. SAI scores related as expected to measures of parental identification and religiosity, and were consistently higher for females than for males. SAI scores did not relate to measures of social desirability response set. The SAI was considered to be a reliable and valid measure of subjective evaluative attitudes.

A review and critique of measures designed to assess the strength of moral values, conscience, or superego (Pittel & Mendelsohn, 1966) led to the conclusion that the instruments which have been used for this purpose are subject to one or more methodological or conceptual criticisms which detract from their usefulness in psychological research. Although the authors pointed to the failure of these tests to predict morally relevant behavior in a direct way, the major concern is not with their predictive validity, but rather with their construction and meaningfulness as measures of moral values (cf. Ebel, 1961).

The critique of previous instruments centered about the following issues which, in addition to the need for standardization data, may be used as guidelines for the construction of more meaningful devices.

1. Test items should describe morally relevant behavior occurring within specified social and motivational frames of reference rather than mere statements of the acts themselves.

2. The range of moral behavior covered should be extensive, and should include acts considered to be morally relevant by the population to be studied in addition to those

aspects of conventional morality usually included in previous instruments.

3. Tests should require the evaluation of acts on the basis of subjective attitudes. They should not assess knowledge of societal standards or awareness of conventional behavior; that is, they should not be contaminated by social desirability response set.

4. Scoring of such tests should not be based on assumptions about the correctness or psychological significance of responses, nor should they require qualitative interpretation by the examiner.

5. Moral values should be assessed directly rather than by inference from assumed indicators such as guilt responses, resistance to temptation, or self-reports of characteristic behavior.

6. Parallel forms should be available for evaluation of changes in values over time or as a function of experimental manipulations.

METHOD

Development of the Situational Appraisal Inventory

Considerations of these guidelines led the present authors to construct a new instrument called the Situational Appraisal Inventory (SAI). It was designed to serve as a measure of moral or evaluative attitudes in a variety of content areas and to provide an overall index of severity of conscience or super-

¹Requests for reprints should be sent to Stephen M. Pittel, Department of Psychiatry, Mount Zion Hospital and Medical Center, 1600 Divisadero Street, San Francisco, California 94115.

ego as these functions are defined in attitudinal terms (cf. Beres, 1958; Sears, 1960).

In the body of this paper the term evaluative attitudes rather than the traditional one, moral attitudes, will be used in order to avoid some of the connotative meanings and content limitations which are associated with the latter term. Evaluative attitudes are inferred from judgments of human conduct along a dimension of "rightness-wrongness." Such judgments may be made about any behaviors, regardless of whether or not they fall into the categories of conventional moral concern.

The SAI consists of items describing the behavior of a fictitious character in a number of specific situations, each of which is to be evaluated by *S* in the light of the framework in which the act takes place. Implicit in each item is the assumption that meaningful judgments cannot be made about behaviors in the abstract. For example, there is high agreement that taking of human life is wrong, but in time of war this value is easily altered; likewise, to steal is generally considered wrong, but the legendary virtue of stealing from the rich and giving to the poor has been taught to generations of children.

The items for the inventory were derived from sample items suggested by a large number of University of California undergraduates who were asked to describe situations in which a person violated some moral code under extenuating or justifying circumstances. The use of this procedure for item construction provided us with acts-in-context which were not necessarily restricted to socially defined morality, but which were, nevertheless, morally salient to college students. Furthermore, the provision that the act be justified within the situational context allowed a sampling of both the content and the motivational aspects of morally relevant situations.

From the initial pool of items thus collected, two 100-item forms of the inventory were compiled (SAI-A and SAI-B). A number of items were randomly chosen to be included in both forms, and other items were repeated within a single form to give a preliminary estimate of response consistency. The following are typical items, chosen to indicate the diversity of situations and behaviors included:

1. John finds a wallet containing a few dollars lying on the sidewalk. Since no one is in sight he pockets the money and leaves the wallet where he found it.
2. John is walking past the door to his sister's room when he notices her lying naked on the bed. This sight arouses him and he stays by her door and secretly watches her.
3. When John goes to the Smorgasbord for dinner he always eats more food than he would ordinarily since he doesn't have to pay extra for it.
4. One day John discovers his father giving a merciless beating to his younger sister. He tries to pull his father away and when he is unable to stop him he hits his father over the head with a paperweight.
5. John has heard that Mary has a reputation for being sexually promiscuous. He ask her out on a

date just for the purpose of having sexual relations with her.

6. John is taking an examination for which he is well prepared. He glances around him and notices that the other students are copying directly from their notes. He realizes that he will not be able to do as well as those who are cheating and so he opens his own notebook and begins to copy from his notes.

The range of content is extensive, but given the source of the items, certain themes predominate, notably sexual behavior, academic honesty, prejudice, authority relationships, and interpersonal honesty. However, an effort was made to transcend these limitations by including items removed from the immediate concerns of a college population, for example, stealing, aggression, job behavior, marital behavior.

For each item, *S* is asked to rate the degree of "wrongness" on an 8-point scale ranging from "Not Wrong" to "Extremely Wrong." The *S*'s ratings are considered a measure of the intensity of evaluative attitudes with respect to the particular situation described. A total score for the SAI is obtained for each *S* by summing the responses to each of the items in a test form. This score may be thought of as a measure of the overall strength of evaluative attitudes independent of the limitations of specific content. As such, it represents a tendency to respond with a characteristic intensity in evaluating human conduct.

In the initial standardization SAI-A was administered to 143 undergraduates (72 males, 71 females), 50 of whom (24 males, 26 females) were retested on the same instrument following a 2-week interval. SAI-B was administered to a similar sample of 209 undergraduates (88 males, 121 females), 79 of whom (35 males, 44 females) were retested on the same form. The *S*s were volunteers from undergraduate psychology courses and included both psychology and nonpsychology majors of all class standings. The mean age for the first sample (UC 1) was 20.46 ($SD = 2.65$) and for the second sample (UC 2) was 19.51 ($SD = 2.87$), a difference significant beyond the .01 level. This difference is probably accounted for by the selection of *S*s in UC 1 largely from a sophomore psychology class, while many UC 2 *S*s were taken from the introductory course.

RESULTS

Reliability of the SAI

Internal consistency. The split-half (odd-even) reliabilities, corrected for attenuation by the Spearman-Brown prophecy formula, are .96 for SAI-A and .97 for SAI-B. Eleven items on Form A and eight items on Form B appeared twice in the same form. For Form A the correlations between first and second ratings of the same item range from .76 to .92 with a median correlation of .87. Since

TABLE 1

TWO-WEEK TEST-RETEST CORRELATIONS, MEANS,
AND STANDARD DEVIATIONS FOR
SAI-A AND SAI-B

Form	N	Test		Retest		r_{1-2}
		\bar{X}	SD	\bar{X}	SD	
SAI-A						
Males	24	422.08	78.67	438.21	82.94	.939
Females	26	465.06	58.94	444.27	59.67	.909
Total	50	443.42	72.61	441.36	71.10	.920
SAI-B						
Males	35	376.50	92.04	368.29	102.62	.951
Females	44	434.91	91.13	407.86	113.69	.955
Total	79	415.11	95.18	390.33	110.04	.952

the instructions on both forms cautioned Ss against turning back or skipping pages, these figures suggest the high stability of item responses to identical situations within a single administration.

Test-retest. Test-retest correlations, means, and standard deviations for total score are given for each sex and for the total sample in Table 1. Reliability coefficients for both sexes and the combined sample all exceed .90.

Frequency distributions of 2-week item reliabilities for SAI-A for each sex and for the combined sample show individual item reliabilities to be only slightly lower than total score reliability. The median reliabilities of .77 for males, .71 for females, and .73 for the combined sample all indicate the high stability of item responses over time. Since Ss were not prepared for a retest in advance

and were specifically instructed "to answer each item as you feel about it now. Do not be concerned with the responses you gave last time," it would appear that the SAI taps a particularly stable and consistent dimension of moral evaluation.

Comparability of SAI-A and SAI-B

The mean total score for SAI-A is 436.90 ($SD = 77.50$), and for SAI-B is 418.43 ($SD = 94.15$) a difference significant at the .05 level ($t = 2.01$). This represents an average item difference between the two forms of only .20, which, considering the 8-point scale used, is quite small. Forty-three females were given Form B followed, after 2 weeks, by Form A. The total score correlation for the two forms is .89. This correlation is only slightly lower than the test-retest correlations for the same form over the same time interval.

Dimensions of Evaluative Attitudes

While the SAI was designed to measure evaluative attitudes in a variety of content areas, no categorization of items according to any preconceived scheme was attempted. Rather, the problem of content scoring was approached through the technique of cluster analysis.

Intercorrelation matrices were obtained for ratings of the 89 nonrepeated items of SAI-A, for the total sample and for the sexes, separately. Tryon key cluster analyses of these three matrices yielded seven cluster dimensions for the combined sample and five for each sex group. The proportion of the total

TABLE 2
DESCRIPTION OF CLUSTERS OF SAI-A: COMBINED ANALYSIS

Cluster	Content	Proportion of communality exhausted	No. of defining items	No. of items with loadings above .50	Mean rating of items loading above .50	Mean correlation of items with total score
1	General	.45	9	18	5.20	.54
2	Sex-Private	.14	4	3	2.53	.35
3	Sex-Illicit	.07	2	3	4.26	.52
4	Prejudice	.12	3	3	4.94	.30
5	Coercion	.06	2	2	6.61	.48
6	Aggression	.08	2	2	2.86	.32
7	Personal Well-Being	.08	2	2	3.31	.39

Note.—N = 72 males and 71 females.

initial communality exhausted by the obtained clusters was .78 for the combined sample, .72 for males, and .77 for females. Table 2 presents data descriptive of the clusters obtained for the combined sample.

The first cluster, which accounts for a large part of the communality, is defined by nine items with a wide diversity of content, ranging from academic honesty to petty larceny to sibling hostility. While few of these items deal with legal violations, all of them are rated as being quite wrong. Further, the 18 items with loadings above .50 on this cluster are characterized by having high correlations with total score on SAI-A. Thus, it would seem that this may best be described as a *general* cluster; an assumption which is supported by the uniform positive loadings of all items on the cluster and by the finding that 58% of the items have their highest loading on this cluster.

Although the remaining clusters each account for only small proportions of the communality they are each marked by the homogeneity of their content and by their psychological meaningfulness. Clusters 2 and 3 deal with sexual behavior. The first of these contains items dealing with masturbation and voyeurism while the latter deals with premarital and extramarital sexual relations. Clusters 3 through 7 are defined by only two or three items and deal, respectively, with anti-Negro prejudice, coercion by force or blackmail, physical aggression, and behaviors which are either directly or indirectly injurious to one's personal well-being.

As shown in Table 3, the intercorrelations among the oblique cluster dimensions are, on the whole, relatively low. As might be expected, the highest correlations are those with the general cluster and those between the two sexual clusters. However, the intercorrelations of items within each cluster (shown in the diagonal) are generally higher than the average intercluster correlation of .29.

Findings similar to the above were obtained in the separate male and female sex analyses. Notable exceptions to these marked similarities include the appearance of a cluster for males centering on petty violations which are almost universally condoned (such as

TABLE 3
INTERCORRELATIONS OF OBLIQUE CLUSTERS
OF SAI-A: COMBINED SEX ANALYSIS

Cluster	1	2	3	4	5	6	7
1	.44						
2	.39	.37					
3	.48	.48	.43				
4	.26	-.05	.11	.40			
5	.52	.28	.40	.37	.38		
6	.31	.40	.16	.02	.29	.29	
7	.48	.18	.25	.01	.31	.34	.33

Note.—Diagonal values are the mean intercorrelations among key cluster variables for each of the seven clusters.

overeating at a Smorgasbord), and the absence of a prejudice cluster in the female analysis. The female analysis also revealed a cluster dealing primarily with interpersonal honesty and a less marked differentiation between the two sexual clusters obtained in both the combined and the male analysis.

On the basis of these cluster analyses 30 SAI-A items were selected for a short form of the inventory (SAI-J). Twenty-three items were selected for their ability to represent a given cluster in a pure fashion, that is, each item had to have a high cluster loading on a single cluster and low loadings on all others. In most cases the items selected by this criterion were the defining, or key cluster, variables for a given cluster. An additional seven items were selected on the basis of their correlation with total score on SAI-A. Items which correlated highly with total score but did not load heavily on any particular cluster were thus chosen for inclusion in Form J. The 30-item SAI-J was constructed in such a way as to ensure that the information available in Form A, in terms of both cluster structure and total score, could be derived more quickly and more efficiently.

Comparability of SAI-A and SAI-J

Although data are not available for a direct comparison of these two forms there is evidence to indicate their comparability. For the combined UC 1 sample ($N = 143$) total score on SAI-A correlated .96 with the score obtained by summing the 30 items which comprise SAI-J. Similarly, the mean ratings for the 30 selected items taken from

the original standardization data correlate .94 with mean ratings for SAI-J items obtained from a sample of 88 male and female Ss who were given only the short form.

Despite these high correlations the ratings of items in SAI-J tend to be somewhat lower than ratings for the same items included in SAI-A. While few of these differences reach statistical significance it is likely that intra-test context does affect item ratings in a consistent fashion. However, the magnitude of these differences is slight and the correlations of item means and total score are sufficiently high to justify the use of the two measures as equivalent.

Validity of the SAI

There are serious difficulties in discussing the validation of an instrument, like the SAI, which is designed to assess subjectively held attitudes. Previous research findings indicate that, with few exceptions, attitude measures are, by themselves, poor predictors of overt behavior. Nevertheless a test which purports to measure evaluative attitudes should have face and construct validity and should differentiate between groups with known value orientations.

Face Validity

Since the SAI was designed as a measure of the strength of evaluative attitudes (i.e., the attitudinal component of conscience or superego), its face validity may be estimated by considering the representativeness of its item content. As indicated above, the items do cover an extensive range of behaviors considered to be morally relevant by college students and the dimensions obtained in the cluster analysis of SAI-A correspond to many of the rubrics of conventional morality. Certainly, the cluster dimensions include the types of behaviors tapped by other measures of this sort in addition to those which reflect the special concerns of college students. Thus, in terms of face validity, the SAI does provide a measure of individual differences in the domain it purports to assess.

Sex Differences in SAI Ratings

Previous research (e.g., Bronfenbrenner, 1961; Rettig & Pasamanick, 1959) indicates

that females show stronger evaluative attitudes than males. The expectation that this sex difference would obtain on the SAI was supported for all forms of the test.

The tendency for females to be more severe in evaluative ratings than males is reflected by their generally higher mean ratings on most items of the inventory although only one-third of these differences reach statistical significance. Those items which do significantly discriminate between the sexes are frequently of a sexual nature or are at least concerned in some fashion with male-female relationships. On Form A of the SAI females have significantly higher ratings on 33 of the 89 nonrepeated items, 19 of which deal with male-female relationships. Females, for example, consider it more wrong for John to marry a girl for her parents' money, to have intercourse with a prostitute prior to his wedding, to neck with his sister, to buy clothes for himself while depriving his wife, or to date a girl specifically with the intention of having sexual relations with her. Females also consider certain nonsexual behaviors to be more wrong than do males. Included in these are items dealing with not donating money to charity, cheating on an examination, use of physical aggression, and attempting to evade the draft.

Although no differences in the opposite direction were statistically significant, males gave higher ratings of wrongness to items dealing with lying to parents, stealing office supplies, arguing with mother, and habitually smoking despite the potential danger to health.

Table 4 presents mean total scores for males and females on the three SAI forms. While females in all samples score significantly higher on total score for all forms of the test, the mean differences across items are sufficiently consistent to result in high agreement about the relative wrongness of items. For SAI-A the correlation between male and female ratings for the 89 nonrepeated items is .93. For SAI-B the two sets of ratings correlate .91, and for SAI-J the correlation is .94.

TABLE 4

COMPARISON OF SAI TOTAL SCORES
OF MALES AND FEMALES

Form	Males			Females			<i>t</i>
	<i>N</i>	\bar{X}	<i>SD</i>	<i>N</i>	\bar{X}	<i>SD</i>	
SAI-A	72	422.08	78.67	71	465.06	58.94	3.69**
SAI-B	88	395.05	91.88	121	435.44	92.87	3.12**
SAI-J	100	119.26	26.96	100	128.00	28.60	2.10*

* $p < .05$.** $p < .01$.*Religious Differences in SAI Ratings*

While the strength of moral (evaluative) attitudes should not necessarily differ as a function of specific religious affiliations, Ss who are strongly religious (regardless of denomination) should have higher SAI scores than those who do not believe or participate in religious matters (cf. Scott, 1965). A comparison of Catholic, Protestant, and Jewish Ss yielded no significant differences in total score for any of the SAI forms (all F s < 1.00). However, with all forms, Ss who claimed *any* religious affiliation obtained higher SAI scores than those who had *no* religious affiliation. Table 5 presents the mean SAI scores for religious and nonreligious affiliated Ss. The relatively small number of Ss in the nonreligious group for SAI-A precluded finding a significant difference, but the data are clearly indicative of the trend demonstrated in each of the other analyses.

The 59 Ss used in the above analysis of religious differences on SAI-A also responded to a questionnaire dealing with the nature and strength of belief systems pertaining to

TABLE 5

COMPARISON OF SAI SCORES: COMBINED RELIGIOUS
AFFILIATIONS AND NONAFFILIATED SUBJECTS

SAI	Combined religious groups			No stated religion			<i>t</i>
	<i>N</i>	\bar{X}	<i>SD</i>	<i>N</i>	\bar{X}	<i>SD</i>	
A	50	463.10	60.83	9	409.88	80.51	1.79*
B	191	431.44	83.79	64	371.94	106.10	4.06**
J	69	121.51	28.74	20	101.75	28.03	2.70**

* $p < .10$.** $p < .01$.

TABLE 6

COMPARISON OF SAI-A SCORES IN TERMS OF RELIGIOUS
BELIEFS AND OBSERVANCE

Religious belief and observance	SAI-A			
	<i>N</i>	\bar{X}	<i>SD</i>	<i>t</i>
Church attendance				
High	18	479.94	48.00	2.22*
Low	41	444.02	70.95	
Strength of belief in God				
High	43	472.51	50.08	2.85**
Low	16	407.88	82.41	
Frequency of thoughts of God				
High	38	477.18	47.95	3.02**
Low	21	418.50	77.36	
Thoughts of God influence Ss behavior				
High	27	478.88	48.18	2.47*
Low	32	437.84	75.22	

* $p < .05$.** $p < .01$.

religion. The four questionnaire items which tapped strength of religious belief and observance are as follows: (a) I attend church: (every week, regularly, only on special occasions, never); (b) I believe that God: (definitely exists, probably exists, probably does not exist, definitely does not exist); (c) Thoughts of God enter my mind: (very often, frequently, occasionally, never); (d) Thoughts of God influence my behavior: (very often, frequently, occasionally, never).

Since relatively few Ss gave extreme answers to any of these items, the data from the first two and second two response categories for each item were combined for analysis. Table 6 presents the mean SAI-A scores for Ss falling into these response categories.

For each of the four items which tap degree of religious belief and observance, those Ss who indicated stronger religiosity obtained significantly higher SAI-A scores than those who were low on religiosity.

In conjunction with the finding that nonreligious affiliated Ss score lower on SAI than those who do state a religious affiliation, it would appear that the strength of religious beliefs is definitely related to the severity of moral evaluation. Thus, the SAI clearly taps a dimension of subjective evaluation which

relates strongly to socially valued belief systems.

Parental Identification and SAI Ratings

Following most other writers in the field (e.g., Sears, 1960), it is assumed that the development of conscience is accomplished through the process of identification. Since it is generally believed that the parents of typical college students represent and transmit to their children the conventional values of their society, it follows that strength of parental identification should be related to the strength of conscience. To the extent that it taps conventional moral (evaluative) attitudes, the SAI should provide an indication of the degree to which Ss have adopted parental values. The results of three studies with the SAI bear on its validity in terms of this hypothesis.

In evaluating these results, it must be kept in mind that each of the measures of identification is crude and open to methodological criticism. Consequently, it is the convergence of the results of these studies that provides evidence of the validity of the SAI.

The Ss who took SAI-A were asked to estimate the number of items on which the judgments of their parents would agree with their own. The estimates of parental agreement was taken as a rough indication of the strength of parental identification. Table 7 presents the mean SAI-A scores of Ss who estimated high agreement with parents and those who indicate low agreement with parents (75% or more vs. 50% or less of the items).

The results are in accord with theoretical expectations. For both sexes, those Ss who estimate high agreement with parents obtain higher SAI scores than those who do not. The difference for males is small and insignificant but the effect is a powerful one for females ($t = 4.95$, $p < .001$).

In a second study Ss were asked to respond to each item on SAI-J in terms of their own judgment, and as they anticipated that their parents would respond. A measure of assumed similarity to each parent based on profile congruence (Bronfenbrenner, 1958) was derived by summing the absolute differences between each S's self- and father ratings

TABLE 7

COMPARISON OF SAI-A SCORES OF GROUPS DIFFERING IN STRENGTH OF PARENTAL IDENTIFICATION (PERCENTAGE OF AGREEMENT)

Identification with both parents	Males			Females		
	N	\bar{X}	SD	N	\bar{X}	SD
High	18	438.89	58.90	18	509.05	38.71
Low	15	422.07	88.82	18	439.50	43.07

and self- and mother ratings across all 30 items. The Ss were classified as being either high or low on identification with each parent by dichotomizing the distribution of assumed similarity scores for each parent separately for each sex. High difference scores indicate weak identification while low difference scores indicate strong identification. Table 8 presents mean scores on SAI-J for groups of Ss who show strong identification with both parents and for those who show weak identification with both parents.

As in the previous study, higher SAI scores are associated with strong identification with both parents. Comparisons of SAI scores for Ss who are not identified with either parent and those who are identified with both parents again yield a significant difference only for females (for males, $t = 1.19$, ns ; for females, $t = 2.70$, $p < .02$). While the data for males are in the expected direction, the difference was not sufficient to achieve statistical significance. It should be noted that the results of the two studies are remarkably similar even though two different estimates of

TABLE 8

COMPARISON OF SAI-J SCORES OF GROUPS DIFFERING IN STRENGTH OF PARENTAL IDENTIFICATION (ASSUMED SIMILARITY)

Identification with both parents	Males			Females		
	N	\bar{X}	SD	N	\bar{X}	SD
High	19	138.45	29.14	16	151.87	25.74
Low	15	126.13	28.97	15	122.00	33.04

strength of identification and two different forms of the SAI were used.

In a third study Distler (1965) derived parental identification scores from a number of questionnaire items dealing with cognitive, affective, and behavioral similarity to parents. Distler was concerned with comparisons between Ss who were primarily mother identified and those who were primarily father identified. The Ss who were not identified with either parent or who were equally identified with both parents were excluded from his analyses. For both sexes, higher SAI-J scores were found for mother identified than for father identified Ss. The ordering of SAI scores in terms of parental identification was females-mother identified (142.04), females-father identified (130.68), males-mother identified (126.30), males-father identified (119.86). For females only, the differences between mother and father identified Ss was significant at the .05 level ($t = 2.09$).

It is clear that SAI scores are related to measures of parental identification in a manner consistent with expectations. High SAI scores are related to strong parental identification for Ss of both sexes and, consistent with the notion that females have more severe evaluative standards, higher SAI scores are associated more with maternal than with paternal identification.

Social Desirability Response Set

SAI total score has been correlated with a number of measures of social desirability response set. For a sample of 100 male and 100 female Ss, the correlations of SAI-J with the Edwards Social Desirability scale (Edwards, 1957) were, respectively, .06 and .15. For these same Ss the correlations of SAI-J and the Marlowe-Crowne measure of social desirability (Crowne & Marlowe, 1960) were .34 for males and .17 for females. Only the correlation of SAI-J and the Marlowe-Crowne scale for males is statistically significant ($p < .001$). Form B of the SAI was administered to a sample of 40 male and female college students who also took the California Psychological Inventory (CPI; Gough, 1957). The correlation of SAI-B with the Communality scale (Cm) of the CPI was .04 for the combined sample. The correlation of

SAI-B with the CPI Good Impression (Gi) scale was .03. These scales are described as measures of "the degree to which an individual's reactions and responses correspond to the modal ('common') pattern established for the inventory" (Cm) and the degree to which a person is "capable of creating a favorable impression . . ." and is "concerned about how others react to them" (Gi) (Gough, 1957, pp. 12-13).

With the exception of the one significant correlation, for males, between SAI-J and the Marlowe-Crowne scale, the results point to the conclusion that SAI scores are not related to any general tendency to respond to personality tests in a socially desirable or typical fashion. Even the one significant correlation coefficient is quite moderate and considerably lower than those reported for most other personality inventories. It is possible, of course, that Ss could, if motivated to do so, "fake good," but in normal research settings this does not appear to affect test scores to an appreciable degree.

DISCUSSION

The SAI provides a very stable measure of evaluative attitudes. While few studies using other measures of moral attitudes report data bearing on either internal consistency or stability of scores, those which do also report high coefficients. For example, Lincoln and Shields (1931), using the Shields Moral Judgments Examination report odd-even reliabilities in four samples of children and young adults of .62, .81, .83, and .93. Friedenberg and Havighurst (1948) report internal consistency reliabilities in the .90s for young adolescents, and Rettig and Pasamanick (1959, 1960, 1961, 1962), using both the Crissman (1942) inventory and a measure of their own, report Kuder-Richardson reliabilities of .93, .95, .96, and .91 for samples of college students, college alumni, blue collar workers and white collar workers. Woodrow (1926) reports a 3-week test-retest correlation of .79. One year test-retest correlations of .50, .65, .66, and .68 are also reported by Friedenberg and Havighurst (1948). Although reliability data are not available for such techniques as story completion, interviews, and a variety of paper-and-pencil and behav-

ioral tests, it still appears that moral attitudes, even in Ss younger than those in the present study, are extremely stable and consistent.

Though the items of the SAI are unusually complex for a psychological test, the reliabilities are in general somewhat higher than those of other instruments in this area. It does not appear necessary to reduce item content to simple abstractions removed from more realistic behavioral contexts to obtain stable judgments. It is possible that the great inter-S variability of ratings for almost all SAI items is attributable to this complexity. Seventy of the 89 unrepeated items in Form A have a 7-point score range, and 89 of the 92 unrepeated items in Form B also have this great a range. This considerably enhances the utility of the test for studies concerned with individual differences in evaluative attitudes.

Another conclusion to be drawn from the data is that Ss respond to the SAI in large part with a generalized evaluative set. Despite the variety of situations included, a number of results support this argument. The high split-half reliabilities indicate that a single dimension accounts for a great proportion of the test variance. Further, for all forms of the test all items correlate positively with total score. The high interform correlation, likewise, indicates a general evaluative set relatively independent of content as does the large general cluster obtained in the cluster analysis of SAI-A. Thus, the interpretation of SAI total score as a measure of the "overall strength of evaluative attitudes independent of the limitations of specific content" seems justified.

This discussion is not meant to imply that the general evaluative set is the only determinant of response to the items. The cluster analysis also yielded clusters of items which were defined by common content characteristics. While total score represents a good estimate of overall strength of evaluative attitudes, Ss do discriminate among items on the basis of content specific judgments. The content clusters are not, however, independent either of each other or of total score.

The objective in the development of the SAI was to provide an instrument suitable for

the systematic study of evaluative attitudes as characteristics of individuals and of social groups. As has been indicated, it seems to fulfill this function both from a psychometric and psychological standpoint. It has been shown that SAI total score is systematically and predictably related to other relevant variables. Differences in SAI scores as a function of sex, religiosity, and strength of parental identification all support the notion that the SAI measures individual differences related to the attitudinal domain of conscience or superego. These results in conjunction with the high reliabilities lead to the conclusion that SAI measures a generalized evaluative set which may be thought of as a consistent personality trait.

To summarize, the SAI is an extremely stable and internally consistent measure of evaluative attitudes. It is largely unaffected by social desirability response set but does relate to sex, religiosity, and strength of parental identification. The great inter-S variability of the SAI in terms of items, clusters, and total score makes it suitable for the study of individual differences related to conscience as well as for the normative study of contemporary evaluative attitudes. The SAI is also suitable as a dependent variable in studies of attitude change, effects of child-rearing practices, patterns of identification, and of reference group affiliation.

REFERENCES

- BERES, D. Vicissitudes of superego functions and superego precursors in childhood. *Psychoanalytic Study of the Child*, 1958, 13, 324-350.
- BRONFENBRENNER, U. The study of identification through interpersonal perception. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior*. Stanford: Stanford University Press, 1958.
- BRONFENBRENNER, U. The changing American child. In E. Ginzberg (Ed.), *Values and ideals of American youth*. New York: Columbia University Press, 1961.
- CRISMAN, P. Temporal change and sexual difference in moral judgments. *Journal of Social Psychology*, 1942, 16, 29-38.
- CROWNE, D. P., & MARLOWE, D. A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 1960, 24, 349-354.
- DISTLER, L. S. Patterns of parental identification: An examination of three theories. Unpublished

- doctoral dissertation, University of California, Berkeley, 1965.
- EBEL, R. L. Must all tests be valid? *American Psychologist*, 1961, 16, 640-647.
- EDWARDS, A. L. *The social desirability variable in personality assessment*. New York: Dryden, 1957.
- FRIEDENBERG, E. Z., & HAVIGHURST, R. J. An attempt to measure strength of conscience. *Journal of Personality*, 1948, 17, 232-243.
- GOUGE, H. G. *California Psychological Inventory manual*. Palo Alto: Consulting Psychologist's Press, 1957.
- LINCOLN, E. A., & SHIELDS, F. J. An age scale for the measurement of moral judgment. *Journal of Educational Research*, 1931, 23, 193-197.
- PITTEL, S. M., & MENDELSON, G. A. Measurement of moral values: A review and critique. *Psychological Bulletin*, 1966, 66, 22-35.
- RETTIG, S., & PASAMANICK, B. Changes in moral values among college students: A factorial study. *American Sociological Review*, 1959, 24, 856-863.
- RETTIG, S., & PASAMANICK, B. Moral codes of American and foreign academic intellectuals in an American university. *Journal of Social Psychology*, 1960, 51, 229-244.
- RETTIG, S., & PASAMANICK, B. Moral value structure and social class. *Sociometry*, 1961, 24, 21-35.
- RETTIG, S., & PASAMANICK, B. Differences in structure and severity of moral judgments by students and by a community sample: Some observations on ethical relativity. *Ohio Journal of Science*, 1962, 62, 317-325.
- SCOTT, W. A. *Values and organizations*. Chicago: Rand McNally, 1965.
- SEARS, R. R. The growth of conscience. In I. Iscoe & H. W. Stevenson (Eds.), *Personality development in children*. Austin: University of Texas Press, 1960.
- WOODROW, H. A picture-preference character test. *Journal of Educational Psychology*, 1926, 17, 519-531.

(Received July 3, 1968)

EXTREME RESPONSE TENDENCY AS A FUNCTION OF EMOTIONAL ADJUSTMENT AND STIMULUS AMBIGUITY

RUSSELL P. NORMAN¹

College of William and Mary

This study attempted to explore relationships between the tendency to check extreme categories on the semantic differential, emotional adjustment, and ambiguity of stimuli to be rated. One hundred and fifteen general psychology students were administered the Welsh *A* and *R* scales of the MMPI and asked to rate nine stimuli on 21 Semantic Differential scales. Stimuli were chosen to reflect three levels of ambiguity (verbal concepts, TAT pictures, and Rorschach inkblots). Three groups of 10 *Ss* each were formed representing high and low combinations of the *A* and *R* scales. Results suggested that anxiety (high *A*, low *R*) is related to extreme response tendencies and repression (low *A*, high *R*) is related to use of the neutral category. "Adjusted" *Ss* (low *A*, low *R*) tend to make more use of the intermediate categories. The relative tendency of high-*A* *Ss* to make extreme responses was most pronounced when stimuli were most ambiguous. Discussion focused on the relationships between scale-checking behavior, "psychopathology," and meaningfulness of material to be rated.

Considerable attention has been paid recently to the study of response sets or styles and the influence these might have on the validity of various psychological test instruments. As a natural extension of these interests, investigators have come to view the problem of response sets in terms of general personality theory. What was formerly thought of as basically "nuisance variance" in a test instrument is now viewed more in terms of motivational variables, the most frequently discussed being social desirability (Edwards, 1957), need for social approval (Crowne & Marlowe, 1960), and acquiescence (Couch & Keniston, 1960).

Most studies which have attempted to view response sets in the broader context of general personality theory have stressed the use of standard objective personality inventories, particularly the MMPI. The true-false or agree-disagree format quite naturally lends itself to ready demonstration of and controlled experimentation with a variety of response styles. It is possible, however, that this approach to the study of personality could make profitable use of additional test instruments, both standardized and experimental,

and that broader and more comprehensive conclusions about the nature and meaning of response sets could be developed.

In their now famous monograph, Osgood, Suci, and Tannenbaum (1957) discuss the issue of scale-checking styles on the Semantic Differential. The authors cite a number of interesting studies suggesting that various scale-checking styles can differentiate normals and schizophrenics (Bopp, 1955) and high- and low-IQ high school students (Kerrick, 1954). They also discuss some possible relationships between generalized anxiety, intelligence, and scale-checking style.

In a recent publication, Zax, Gardiner, and Lowry (1964) discovered that "maladjusted" *Ss* tended to use extreme categories more than intermediate categories when rating projective material (Rorschach inkblots and Blacky pictures) on the Semantic Differential, while "adjusted" *Ss* used more intermediate and fewer extreme categories. They argued that the immature or poorly adjusted individual was more likely to organize his cognitive field in terms of extreme categories, and ample theoretical basis for such a formulation was derived from such diverse areas as behaviorism (Spence, 1958), psychoanalysis (Fenichel, 1945), and Gestalt psychology (Werner, 1948). The authors suggested that not only

¹Requests for reprints should be sent to Russell P. Norman, Department of Psychology, College of William and Mary, Williamsburg, Virginia 23185.

psychopathology but probably also developmental level influences this tendency to endorse extreme rather than intermediate categories, and that research in this area could add relevant contributions to a growing body of information on the relationship between response style and personality variables.

It occurred to the present writer that additional variables were probably quite relevant to the issue of extreme response tendency, those first coming to mind being the nature of the emotional maladjustment and the type of stimulus being rated on the Semantic Differential. Zax et al. (1964) used the Rorschach (a relatively ambiguous stimulus) and the Blacky (a less ambiguous stimulus). Since the results with the Blacky did not confirm their hypothesis and the Blacky was used only with the children's groups, one might be as easily drawn to a level of ambiguity interpretation as a developmental one.

By the same token if one equates emotional maladjustment with the presence of undefended anxiety, then response extremity on the Semantic Differential might be an understandable occurrence. There is a considerable body of evidence which attests to the fact that anxiety narrows the cognitive field and rigidifies behavior (see a recent review of literature by Korchin, 1964). But what of the maladjusted individual without significant overt or manifest anxiety? Would such a relationship hold for the overcontrolled and repressed person as well? Since these two dimensions of maladjustment have consistently emerged in factor-analytic studies (Adams, 1964; Welsh, 1956), it would seem meaningful to study response extremity in relation to each.

The writer² attempted to follow up the findings of Zax et al. by using a more precise measure of maladjustment (Welsh *A* and *R* scales of the MMPI) and having *Ss* rate four verbal concepts (College Professor, President Johnson, Negro, Myself) rather than the more ambiguous Rorschach stimuli. When high-*A*-low-*R* and high-*R*-low-*A* *Ss* were compared with *Ss* low on both *A* and *R*, no significant

stylistic differences of any kind could be detected on the Semantic Differential. It is possible that the *A* and *R* scales were not adequately measuring emotional maladjustment, or more likely, that the extreme response tendency might not appear in maladjusted *Ss* when more structured verbal concepts are being rated.

In deference to the second possibility, the present experiment was devised to study the relationships between emotional adjustment, stimulus ambiguity, and response style on the Semantic Differential. It was hypothesized that in all *Ss* there would be a tendency to make greater use of the extreme scale positions as the stimulus material increased in ambiguity, and that this tendency would be significantly more pronounced in the maladjusted *Ss*. No specific predictions were made concerning the anxiety-repression dimension, but as a working hypothesis, it was assumed to be more likely that *Ss* high on *R* would fall between *Ss* high on *A* and "adjusted" *Ss* in terms of frequency of extreme responses.

METHOD

The *Ss* were 115 general psychology students at the College of William and Mary. The 79 items of the Welsh *A* and *R* scales were randomly ordered and administered to all *Ss* in a single group session. The questionnaires were scored and three groups were formed based on high and low scores on the two scales. The three groups and their *A* and *R* means were as follows: High *A* (24.3), Low *R* (12.2); Low *A* (7.5), High *R* (21.1); and Low *A* (7.8), Low *R* (13.2). A fourth possible group (High *A*, High *R*) was excluded since only three individuals would have qualified for inclusion. Total *N* was 30, with 10 in each group.

In a second group session, *Ss* were asked to rate nine stimuli on 21 Semantic Differential scales, which were the same as those used by Zax et al.³ The stimuli represented three groupings along the dimension of stimulus ambiguity. Three verbally stated concepts (College Professor, President Johnson, Negro) were assumed to represent the least degree of stimulus ambiguity. Three pictures from TAT cards (3BM, 8GF, and 17BM) were chosen to represent moderately ambiguous stimuli, and three Rorschach

³ Tense-Relaxed; Ugly-Beautiful; Hot-Cold; Cautious-Reckless; Large-Small; Wise-Foolish; Happy-Sad; Rounded-Angular; Brave-Cowardly; Sharp-Dull; Good-Bad; Fair-Unfair; Hard-Soft; Rough-Smooth; Kind-Cruel; Clean-Dirty; Fast-Slow; Strong-Weak; Active-Passive; Masculine-Feminine; Heavy-Light.

² R. P. Norman. Emotional adjustment and response extremity: The rating of verbal concepts. Unpublished manuscript, 1965.

cards (I, V, VII) were selected as highly ambiguous stimuli. The nine stimuli were presented to the entire group in balanced sequence (ABCABCABC) by means of an opaque projector. No special criteria were used to select the stimuli within each ambiguity grouping. It was felt, however, that the verbal concepts should be salient to college students, and that the TAT stimulus material should be clear enough in detail to be noticeably less ambiguous than the Rorschach cards.

The mimeographed scales of the Semantic Differential were stapled together in such a way that three random orderings would occur in equal numbers for each stimulus, and no *S* would be exposed to the same ordering for consecutive stimuli. Adjective position was also varied so that no one adjective appeared on the same side of the scale on all three orderings. Semantic Differential scoring consisted of a simple frequency count of the use of the three primary positions for each *S*; extreme responses (two scale positions), neutral responses (one scale position), and intermediate responses (four scale positions).

RESULTS

The mean frequencies for extreme, neutral, and intermediate responses in these three groups are shown in Tables 1-3. A two-way analysis of variance was carried out for each response category.

Extreme Responses

The three groups are significantly differentiated ($F = 3.75$, $df = 2/81$, $p < .05$) with the "adjusted" *Ss* making extreme responses less frequently in each stimulus condition. In addition, the anxious *Ss* make extreme responses most frequently, and consistent with the working hypothesis, the "repressed" *Ss* are intermediate between the other two groups. In terms of the stimulus dimension, the TAT seems to elicit more extreme responses than do the verbal concepts or the Rorschach. This is true to some extent in each experimental

TABLE 1

MEAN NUMBER OF EXTREME RESPONSES IN THREE
EXPERIMENTAL GROUPS FOR THREE
DIFFERENT STIMULI

Group	Verbal	TAT	Ror- schach	<i>M</i>
High <i>A</i> , Low <i>R</i>	8.0	12.7	10.8	10.5
High <i>R</i> , Low <i>A</i>	9.1	11.3	5.8	8.7
Low <i>A</i> , Low <i>R</i>	5.2	8.5	2.9	5.5
<i>M</i>	7.4	10.8	6.5	

TABLE 2

MEAN NUMBER OF NEUTRAL RESPONSES IN THREE
EXPERIMENTAL GROUPS FOR THREE
DIFFERENT STIMULI

Group	Verbal	TAT	Ror- schach	<i>M</i>
High <i>A</i> , Low <i>R</i>	18.2	14.3	19.3	17.3
High <i>R</i> , Low <i>A</i>	20.8	19.4	27.6	22.6
Low <i>A</i> , Low <i>R</i>	13.4	11.5	21.1	15.3
<i>M</i>	17.5	15.1	22.7	

group, although the *F* ratio ($F = 3.05$, $df = 2/81$) barely fails to attain significance at the .05 level. Simple analyses of variance (Stimuli \times Groups) were run, and only the Rorschach significantly differentiated the three groups ($F = 3.35$, $df = 2/27$, $p = .05$). With respect to extreme responses, then, the ambiguity hypothesis receives some support since the most ambiguous stimuli are the only ones to differentiate significantly between the groups. The Rorschach findings are also consistent with those of Zax et al., although a somewhat different picture emerges when the stimulus dimension is changed in the direction of lessened ambiguity.

Neutral Responses

Analysis of neutral responses reveals that both groups ($F = 3.84$, $df = 2/81$) and stimuli ($F = 4.10$, $df = 2/81$) are differentiated at the .05 level. Apparently, the ambiguity of the Rorschach blots selected for this study resulted in more neutral responses than either of the other two stimulus categories. Regardless of condition, however, the repressed *Ss* demonstrate a greater tendency to make neutral responses. On the verbal concepts and the

TABLE 3

MEAN NUMBER OF INTERMEDIATE RESPONSES IN THREE
EXPERIMENTAL GROUPS FOR THREE
DIFFERENT STIMULI

Group	Verbal	TAT	Ror- schach	<i>M</i>
High <i>A</i> , Low <i>R</i>	37.8	36.0	32.9	35.5
High <i>R</i> , Low <i>A</i>	34.1	32.3	29.6	32.0
Low <i>A</i> , Low <i>R</i>	44.5	43.0	38.0	41.8
<i>M</i>	38.8	37.1	33.5	

TAT, the anxious Ss are intermediate between the other two groups, although this does not occur on the Rorschach.

Intermediate Responses

Intermediate responses significantly distinguish between the three groups beyond the .01 level ($F = 7.02$, $df = 2/81$). Regardless of stimuli, the "adjusted" group makes a greater use of this type of response. There is a slight and statistically insignificant tendency ($F = 2.07$, $df = 2/81$) for intermediate responses to decrease with increasing ambiguity of the stimulus.

DISCUSSION

The results of the present study seem to suggest that anxiety is related to extreme response tendencies, repression to neutrality of response, and that "normal" Ss seem to avoid both of these tendencies and stress the use of intermediate categories. The data also suggest that the nature of the stimulus is related to these response tendencies but not necessarily in terms of ambiguity. Verbal stimuli may be less ambiguous than projective test material, but it is obvious that they differ in other ways as well. If the verbal concepts are more salient, it may well negate the relevance of certain Semantic Differential scales, thus forcing more responses into the neutral category. For similar reasons, scales that are more relevant to a particular concept may result in a greater use of extreme categories.

Of the three stimulus conditions, the TAT elicited more extreme responses, while the Rorschach was associated with more neutral responses. It could be that the human content depicted by the TAT pictures was personally more meaningful than the relatively vague and amorphous inkblots. If this were true, it would lend support to the theory which states that response extremity can be viewed not only as a function of psychopathology but also as a function of the meaningfulness of the material to be rated and the rating categories themselves (see a recent review of the literature on response extremity by O'Donovan, 1965). The fact that the Rorschach elicited more neutral responses might be a result of the particular blots selected, but could also be explained by the fact that less ambiguous

stimuli were present in the same rating task. In other words, after viewing and rating material that was more salient and/or more meaningful, it is possible that Ss perceived many of the 21 scales as irrelevant to the Rorschach blots. In fact, several Ss placed nearly all of their Rorschach ratings in the neutral category. It is clear that in future research in this area, it will be necessary to define stimulus categories more precisely and take into account the dimension of meaningfulness in terms of the stimuli and rating scales.

In discussing the findings in relation to the Welsh *A* and *R* scales, certain issues must be resolved in this regard as well. The writer does not feel it is relevant to use the term "psychopathology" when studying an essentially normal population. Since the *A* and *R* scales have been developed from self-descriptive statements, one may infer that high-*A* sensitizers *describe themselves* in somewhat deviant, socially objectionable terms, while high-*R* repressors describe themselves in a conventionally acceptable manner. This means that the status of these individuals in terms of psychopathology must be assessed by means of additional criteria. Clinicians working with college students are familiar with those individuals who show very deviant MMPI profiles, but who fail behaviorally to show a degree of disturbance commensurate with their self-descriptions. Technically, these might be called "false positives," but their existence demonstrates the futility of attempting to determine degree of psychopathology purely in terms of the symptoms a person is willing to admit or deny.

Other evidence reported by Ourth (1963; cited in O'Donovan 1965) suggests that individuals who use extreme rating categories are more capable of showing improvement in psychotherapy than those who tend toward the use of neutral categories. It is generally assumed that the best candidates for psychotherapy are subjectively uncomfortable (admit symptoms), personally sensitive to a wide range of stimuli, and willing to view themselves and their difficulties from an internal frame of reference. These might well be the same individuals who have been referred to as

"sensitizers." Conversely, those who are perhaps least likely to profit from psychotherapy tend to deny subjective discomfort, are personally sensitive to only a narrow range of stimuli, and tend to view their problems from an external frame of reference, a description which closely resembles that of the "repressor" category.

The results of this study suggest that stimulus ambiguity might tend to maximize the basic tendencies of these two groups. One might expect the sensitizers to see the stimuli as more meaningful and give them subjective interpretations, while the repressors will tend to move in the opposite direction and see the stimuli as less meaningful or irrelevant. This, of course, is related to the rationale of most projective tests, and would be consistent with the observation that anxious individuals give more varied and dynamically significant responses, while repressed individuals are typically unproductive and defensive.

Thus, the writer does not find it difficult to reconcile the psychopathology and meaningfulness hypotheses at least in relation to a nonpsychiatric population. It must also be remembered that group averages tend to obscure individual data and exaggerate the differences that do exist. All Ss in the present study were drawn from a relatively homogeneous population in terms of age, ability, socioeconomic background, and psychiatric status. Correlational analysis of these data for all Ss would no doubt reveal positive but uniformly low coefficients. One is forced to agree with O'Donovan that future research in this area should focus on the interaction between stimulus and subject characteristics. In addition, the writer feels that it will be mandatory to move beyond such overly simplified concepts as psychopathology, as well as exclusive reliance on paper-and-pencil instruments.

REFERENCES

- ADAMS, H. B. "Mental illness" or interpersonal behavior? *American Psychologist*, 1964, 19, 191-197.
- BOFF, J. A quantitative semantic analysis of word association in schizophrenia. Unpublished doctoral dissertation, University of Illinois, 1955.
- COUCH, A., & KENISTON, K. Yeasayers and naysayers: Agreeing response set as a personality variable. *Journal of Abnormal and Social Psychology*, 1960, 60, 151-174.
- CROWNE, D. P., & MARLOWE, D. A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 1960, 24, 349-354.
- EDWARDS, A. L. *The social desirability variable in personality assessment and research*. New York: Dryden, 1957.
- FENICHEL, O. *The psychoanalytic theory of neurosis*. New York: Norton, 1945.
- KERRICK, J. S. The effects of intelligence and manifest anxiety on attitude change through communications. Unpublished doctoral dissertation, University of Illinois, 1954.
- KORCHIN, S. J. Anxiety and cognition. In C. Scheerer (Ed.), *Cognition: Theory, research, promise*. New York: Harper, 1964.
- O'DONOVAN, D. Rating extremity: Pathology or meaningfulness? *Psychological Review*, 1965, 72, 358-372.
- OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning*. Urbana: University of Illinois Press, 1957.
- OURTE, L. L. The relationship of similarity in therapist-client pairs to client's stay and improvement in psychotherapy. Unpublished doctoral dissertation, University of Missouri, 1963.
- SPENCE, K. W. Behavior theory and selective learning. *Nebraska Symposium on Motivation*, 1958, 6, 73-107.
- WELSH, G. S. Factor dimensions A and R. In G. S. Welsh & W. G. Dahlstrom (Eds.), *Basic readings on the MMPI in psychology and medicine*. Minneapolis: University of Minnesota Press, 1956.
- WERNER, H. *Comparative psychology of mental development*. New York: Follett, 1948.
- ZAX, M., GARDINER, D. H., & LOWRY, D. G. Extreme response tendency as a function of emotional adjustment. *Journal of Abnormal and Social Psychology*, 1964, 69, 654-657.

(Received July 8, 1968)

PERSONALITY FACTORS RELATED TO RESPONSE TO SOCIAL ISOLATION AND CONFINEMENT¹

DALMAS A. TAYLOR,² IRWIN ALTMAN, LADD WHEELER³
AND ESTELLE N. KUSHNER

Naval Medical Research Institute, Bethesda, Maryland

The present study examines (a) personality factors relating to or predictive of endurance and adjustment to social isolation, and (b) personality and/or perceptual changes occurring under varying conditions of social isolation. Pairs of men were confined for 8 days in small rooms without recreational materials of any kind and a minimal work schedule. Multiple conditions of isolation were effected by a factorial combination of mission-length expectation, stimulation, and privacy. Eight Holtzman Inkblot Technique variables showed significant changes from before to after confinement. Most of these changes were among variables included in Factors I and II (Holtzman et al., 1961). Movement, Color, Penetration, and Barrier changes were associated with varying conditions of confinement. Interpretation of Factor I changes was confounded by corresponding changes in verbal productivity. Other instruments used provided only marginal findings as regards prediction or adjustment. A post hoc analysis of disruptive groups yielded a syndrome of adjustment possibly related to extreme or severe conditions of social isolation. Suggestions for future research leading to selection and training were discussed.

Investigations of personality factors associated with endurance in prolonged sensory and social deprivation (Hull & Zubek, 1962; Myers, 1969; Wright & Zubek, 1966) and with psychological adjustment to isolated environments (Cole, Machir, Altman, Haythorn, & Wagner, 1967; Fulkerson, Freud, & Raynor, 1958; Sells, 1961; Wright, Chylinski, Sisler, & Quarrington, 1967; Wright, Sisler, & Chylinski, 1963) have achieved only limited success in discriminating between well-adjusted and poorly adjusted individuals.

In a review of research, Myers (1969) concluded that personality patterns relating to endurance of individuals in prolonged

sensory deprivation (SD) situations can be distinguished only through "combinational equations predictive of endurance or rough directional trends from many massive *N* experiments [p. 24]." Myers also stated that unsuccessful *Ss* are higher on the Psychopathic Deviate (*Pd*) and Hypomania (*Ma*) scales of the MMPI, and on the n-Aggression scale of the EPPS. Successful *Ss*, on the other hand, are higher on the Consistency, n-Deference, and n-Affiliation scales of the EPPS, and higher on the Social Responsibility (*Re*) and Ego Strength (*Es*) scales of the MMPI. Wright et al. (1967) investigated adjustment in isolated, semiisolated, and nonisolated environments and found poor adjustment associated with high *Pd* and *Ma* scores, and low *Es* scores primarily in semiisolated settings. Additionally, poorly adjusted groups in the isolated and semiisolated settings were lower on n-Exhibitionism and Heterosexuality and higher on n-Intracception on the EPPS than well-adjusted groups.

Projective inkblot tests have also been used with meager results. Wright and Zubek (1966) used the Rorschach and the Buhler-Lefever Sign List in a multiple discriminant function analysis of successful and unsuccessful perceptually deprived *Ss*, but failed to

¹From Bureau of Medicine and Surgery, Navy Department Research Subtask MF022.01.03-1002. The opinions and statements contained herein are the private ones of the writers and are not to be construed as official or reflecting the views of the Navy Department or the Naval Service at large. The authors wish to thank William N. Colson, Ira Donenfeld, Evelyn Lett, Leonard Oberlander, and Richard Sorrentino for their help in the collection and processing of the data. The authors are also indebted to Elizabeth Broomhead, Donald Gorham, and Wayne H. Holtzman for their comments on earlier versions of the manuscript.

²Requests for reprints should be sent to Dalmás A. Taylor, NMRI-032-1b, National Naval Medical Center, Bethesda, Maryland 20014.

³Now at New York University.

replicate an earlier study which distinguished successful and unsuccessful Ss. Cole et al. (1967) found perceptual changes, for isolates only, as measured by the Holtzman Inkblot Technique (HIT), on Human form following 10 days of social isolation, as well as decrements in Location, Form Definiteness, and Movement and an increase in Color for both isolation and control groups. They failed to find any relationship between HIT changes and personality composition factors related to compatibility and incompatibility. Subsequent analyses of this data by the present authors, combining Color and Movement into an M/C ratio, showed incompatible dyads significantly decreasing in M/C ratio from preisolation to postisolation. Cole et al. (1967) interpreted declines in Movement and increases in Color as indicative of Ss being "more sensitive to external stimuli while utilizing internal resources less as the stress of experimental conditions continued [p. 332]." This interpretation is also supported by the subsequent finding of a decrease in the M/C ratio for incompatible groups. Furthermore, postexperimental questionnaires from the same Ss indicated that incompatible groups experienced higher levels of stress than compatible groups (Haythorn, Altman, & Myers, 1966).

The purposes of the present study were to examine further (a) personality factors relating to or predictive of endurance and adjustment to social isolation and (b) personality and/or perceptual changes occurring under varying conditions of social isolation. Specifically, social isolation conditions were manipulated according to (a) *privacy*, determined by whether two men lived in one room or in two adjoining rooms with free access between the rooms, (b) *stimulation*, determined by whether or not groups had various forms of enrichment and contact with the outside world, and (c) *mission-length expectation*, or whether groups expected to be isolated for relatively short periods (4 days) or relatively long periods (20 days). Data reported elsewhere (Taylor, Wheeler, & Altman, 1968) indicated that privacy, stimulation, and short missions yielded least stress reactions, whereas groups expecting long missions, having no outside stimulation and in a privacy condition

were most stressed and debilitated. A key question posed here is whether personality and/or perceptual changes will vary in accordance with these differentially stressful situations.

METHOD

Pairs of men were confined for 8 days in small rooms without recreational materials of any kind and a minimal work schedule. Diet consisted of liquid food, survival crackers, and water. Lights remained on at all times to allow continuous monitoring by closed-circuit television and ceiling microphones.

The Ss were 18- to 20-year-old volunteers who had just completed boot training at the Great Lakes Naval Training Center. Ninety-five percent of 168 of the men requested to volunteer on arrival at the Naval Medical Research Institute did so.⁴ Pairs were selected such that members were not only strangers but roughly equated for age, religion, IQ, education, and other demographic variables.

In half of the groups, Ss had separate rooms at their disposal (privacy condition); the other half lived in one room (no-privacy condition). In all cases, the rooms were similar in size and were furnished identically.

A stimulation manipulation permitted an examination of the impact of loss of contact and separation from cultural experiences. The Ss in the no-stimulation condition did not hear another human voice. Communications were given by tone and buzzer code system. By contrast, Ss in the stimulation condition received verbal instructions to perform all tasks, were asked to report room temperature and food consumption, and received three 5-minute broadcasts per day.

After the isolation period had begun, half of the groups were told that the mission would last 4 days, and the remaining half were informed that the mission would last 20 days. The mission actually lasted 8 days for both groups. Both groups were told that the mission might be extended "due to military operational requirements." At the end of 4 days, no announcement was made to the 4-day groups regarding the extension of their mission. This combination of experimental variables yielded a $2 \times 2 \times 2$ factorial design.

The dependent variable measures used to assess personality and perceptual factors were the Minnesota Multiphasic Personality Inventory (Hathaway & McKinley, 1951), the Edwards Personal Preference Schedule (Edwards, 1959), the Holtzman Inkblot Technique (Holtzman, Thorpe, Swartz, & Herson, 1961), and a sociopathy scale (Lykken, 1957). The

⁴ The authors owe a special debt of thanks to Orren R. Royal, United States Navy, for serving as Medical Officer on the project, administering all preexperimental and postexperimental physical and psychiatric examinations, and making himself available 24 hours a day for almost a year.

TABLE 1

PRE-POST MEAN SCORES FOR ISOLATION
GROUPS ON HIT

Scales	Pre	Post	F
Form Definiteness (<i>FD</i>)	77.41	71.72	9.71*
Color (<i>C</i>)	9.23	12.60	20.69**
Shading (<i>Sh</i>)	3.45	5.06	11.43*
Movement (<i>M</i>)	29.10	23.38	13.41*
Integration (<i>I</i>)	5.33	3.60	43.16**
Human (<i>H</i>)	22.38	15.72	48.00**
Penetration (<i>Pn</i>)	4.58	3.42	10.56*
Location (<i>L</i>)	40.20	48.42	41.33**
Total Words Used (<i>TWU</i>)	168.43	139.92	29.62*

* $p < .01$.** $p < .001$.

MMPI and Lykken scales were administered 1 or 2 days before confinement and as soon after confinement as possible; the EPPS was given pre-confinement only. The HIT was group administered with Forms A and B counterbalanced from pre-confinement to postconfinement.⁵ It was given the evening before isolation began and, in all cases, within an hour or two following termination of the experimental run.

RESULTS

Holtzman Inkblot Technique

Eight of the HIT scales showed significant changes from before to after social isolation (see Table 1). Significant decreases occurred for Form Definiteness (*FD*), Movement (*M*), Integration (*I*), Human (*H*), and Penetration (*Pn*). Significant increases occurred for Location (*L*), indicating use of smaller portions of the blot, Color (*C*), and Shading (*Sh*). Additionally, Total Words Used (*TWU*) decreased from preexperimental to postexperimental testing. These findings replicate and extend those of Cole et al. (1967), the only difference being that the earlier study did not find pre-post changes in *I*, *Pn*, and *Sh*.

Six of the eight HIT determinants showed pre-post changes are included in two of the empirically derived factors identified by Gorham (1967) and Holtzman et al. (1961).

⁵ Forms A and B of the HIT were projected on a screen from 35-mm. color slides produced at the Naval Medical Research Institute with the kind permission of the Psychological Corporation, New York. The authors would like to thank Donald R. Gorham for his computer scoring of the inkblot protocols.

TABLE 2

MEAN NUMBER OF COLOR RESPONSES
BEFORE AND AFTER ISOLATION

Days	Isolation	Pre	Post
4	Private	8.36	10.28
	No Private	7.44	11.25
20	Private	10.67	17.08*
	No Private	10.56	12.61

* Significantly larger than all other means ($p < .01$) by Newman-Keuls test.

Human, Integration, Movement, and Form Definiteness constituted one factor considered to "be indicative of well organized, ideational activity and good imaginative capacity [Holtzman et al., 1961, p. 171]." Postconfinement decrement occurred on all scales in this factor. The other factor included Color and Shading, thought to depict oversensitivity to the stimulus properties of the inkblots. In the present study, Color and Shading increased from preisolation to postisolation. Thus, in agreement with research on sensory deprivation (Bexton, Heron, & Scott, 1954), social isolation seems to have produced an impoverishment of cognitive-perceptual activity.

One interest of the present study was to examine the possibility that these changes might vary according to different conditions of social isolation. Specifically, Movement and Color changes were expected to relate to stress such that simultaneous increases in Color and decreases in Movement would be greater under the most stressful isolation conditions. A significant Mission Length \times Privacy \times Pre-Post interaction occurred for Color ($F = 4.59$, $p < .05$). Examination of means associated with this interaction (see Table 2) indicates that all groups showed increments in Color from preisolation to postisolation, although the increase was significant only for the 20-day privacy group, whose posttest Color scores were also significantly higher than those of all other groups. It is interesting that this condition also yielded the highest subjective reports of stress and anxiety (Taylor et al., 1968). Aborters (i.e., Ss who could not complete the 8 days of confinement) in the privacy condition also showed significant increases in Color ($F =$

10.29, $p < .01$) from preisolation to postisolation and had higher postconfinement means for Color than all other groups.

Further, there were significant Mission Length \times Stimulation \times Privacy \times Pre-Post interactions for Shading ($F = 5.33$, $p < .05$) and Movement ($F = 5.17$, $p < .05$). Newman-Keuls analyses on the means for Movement showed significant decrements for the 4-day stimulation no-privacy groups only. Means for Shading revealed no differential pre-post changes. Perhaps short missions affording enrichment and interpersonal contacts minimized perceptual experiences related to Movement. Similarly, 4-day groups decreased significantly on Penetration and Barrier. No interpretation of these latter two changes is offered. Pre-post decreases on Human and Integration were parallel for aborters and completers, with no postconfinement differences.

In summary, pre-post increases in Color were associated with long mission expectations (20 days), privacy, and unsuccessful mission completion, whereas postconfinement decrements for Movement, Barrier, and Penetration occurred for various combinations of short mission expectations (4 days). The finding regarding Color confirms other data relating Color change to heightened affect (see Cerbus & Nichols, 1963). Lowered Barrier and Penetration scores may relate to better adjustment under less stressful conditions of confinement.

A rather unexpected finding was a decrement in total verbal output for all groups, although they were least for aborter groups and for groups under the most stressful conditions—20-day privacy (see Taylor et al., 1968).

There were no preexperimental abort-completer differences in HIT determinants, indicating the lack of overall utility of the test in predicting adjustment to isolation. However, as will be discussed, there were several preconfinement HIT differences for groups showing extreme responses to the isolation situation.

Minnesota Multiphasic Personality Inventory

There were six MMPI scales which produced pre-post isolation changes, although the

TABLE 3
PRE-POST MEAN SCORES FOR ISOLATION
GROUPS ON MMPI SCALES

Scales	Pre	Post	F
Depression (2)	53.10	49.58	6.35***
Hysteria (3)	52.10	50.48	4.98**
Psychopathic Deviate (4)	57.30	54.43	8.41***
Hypomania (9)	60.08	63.27	8.72***
Introversion (0)	50.18	48.35	7.32***
Anxiety (MA)	12.05	13.21	3.10*

Note.—T scores with K corrections.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

absolute magnitude of change was small (see Table 3). Depression, Hysteria, Psychopathic Deviate, and Social Inversion decreased and Hypomania and Manifest Anxiety increased from preconfinement to postconfinement. Presumably, social isolation had the effect of abating elevations on some of the clinical scales of the MMPI, but did produce elevated states of anxiety and hyperactivity.

Higher order interactions occurred but did not yield any systematic relationship to experimental treatments. For example, several pre-post changes were associated with the stimulation manipulation. However, these were difficult to interpret since Newman-Keuls analyses revealed that stimulation groups had higher preconfinement means for Hysteria, Psychopathic Deviancy, and Psychasthenia and did not differ from the no-stimulation groups on postconfinement assessment. Hence, these differences might be due to sampling problems.

A Privacy \times Pre-Post interaction for Psychopathic Deviancy ($F = 3.98$, $p < .05$) and a significant Stimulation \times Privacy \times Pre-Post interaction for Hysteria ($F = 5.53$, $p < .05$) indicated that no-privacy Ss decreased from preisolation to postisolation on the Psychopathic Deviate scale, and the stimulation-privacy Ss also decreased on the Hysteria scale. However, the preexperimental mean in the latter case was significantly higher than all other preconfinement means associated with the Stimulation \times Privacy interaction.

Analyses of changes in MMPI scores for aborters and completers in relationship to the

TABLE 4

ONE-WAY ANALYSES OF VARIANCE (KRUSKAL-WALLIS) FOR RIOT-ABORTERS, ABORTERS, AND COMPLETERS ON SELECTED PERSONALITY TRAITS

Variable	Pre				Post			
	Riot- abort <i>M</i>	Abort <i>M</i>	Com- plete <i>M</i>	<i>H</i>	Riot- abort <i>M</i>	Abort <i>M</i>	Com- plete <i>M</i>	<i>H</i>
MMPI								
<i>F</i> (validity)	4.0 _a	1.8 _b	6.0 _a	4.30	21.0 _a	6.0 _b	3.2 _b	6.30**
Hypochondriasis (<i>Hs</i>)	49.0 _a	41.2 _a	49.8 _a	4.41	71.2 _a	40.0 _b	44.5 _b	8.09**
Depression (<i>D</i>)	44.5 _a	50.2 _a	49.5 _a	1.24	63.2 _a	47.8 _b	50.2 _b	4.84*
Paranoia (<i>Pa</i>)	50.0 _a	43.2 _a	52.2 _a	1.29	72.5 _a	44.0 _b	46.2 _b	7.48***
Psychasthenia (<i>Pt</i>)	54.5 _a	48.5 _a	48.5 _a	2.44	65.8 _a	46.5 _b	50.0 _b	7.57***
Schizophrenia (<i>Sc</i>)	56.5 _a	51.8 _a	54.8 _a	1.64	91.0 _a	44.0 _b	48.2 _b	7.90***
Repression-Sensitization (<i>R-S</i>)	34.2 _a	25.2 _{ab}	17.5 _b	4.29	46.5 _a	28.0 _{ab}	24.2 _b	6.11***
Control (<i>Cn</i>)	59.5 _a	53.0 _{ab}	45.2 _b	5.76**	58.0 _a	49.2 _a	54.2 _a	1.71
Responsibility (<i>Re</i>)	42.2 _a	46.8 _{ab}	58.5 _b	5.86**	43.0 _a	49.8 _a	53.0 _a	4.84*
EPPS								
Heterosexual	26.0 _a	23.2 _a	19.0 _b	7.72***	—	—	—	—
Aggression	15.8 _a	9.8 _b	9.8 _b	4.24	—	—	—	—
Primary sociopathy								
Lykken	26.2 _a	19.5 _a	20.5 _a	3.90	29.0 _a	17.0 _b	22.5 _b	5.83*
Affect/stress								
Subjective Stress Scale	10.8 _a	46.0 _b	8.8 _a	7.70***	55.2 _{ab}	72.2 _a	15.5 _b	4.67*
Anger (PAS)	12.9 _a	3.7 _b	6.6 _b	4.86**	12.5 _a	18.4 _a	3.7 _b	8.00***
HIT								
Animal (A)	31.8 _a	25.9 _{ab}	21.5 _b	3.46	25.5 _a	31.2 _a	17.0 _b	6.32**
Hostility (Hs)	7.0 _{ab}	3.8 _a	9.0 _b	5.07	3.8 _a	5.2 _a	5.5 _a	.89
Barrier (Br)	7.5 _{ab}	7.0 _a	11.0 _b	4.65*	7.2 _a	8.0 _a	9.2 _a	1.25
Popular (HP)	5.0 _a	5.8 _a	5.8 _a	5.07**	3.5 _a	3.0 _a	4.5 _a	.52

Note.—Items within a row having a single subscript in common do not differ significantly from each other; items not sharing a letter in common are significant at the .05 or .01 level. These comparisons are not appropriate between preconfinement and postconfinement means.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

experimental conditions revealed no meaningful relationships, nor did they indicate any preexperimental differences among Ss.

In general, there were few significant relationships between MMPI preexperimental scores or changes in scores and success in completing the mission or as a function of experimental treatments. However, pre-post changes indicated better adjustment (postconfinement decreases on *D*, *Hy*, *Pd*, and *Si* scales) and increased activity and anxiety (postconfinement increases on the *Ma* and *MA* scales). The absence of preexperimental differences between aborters and completers indicates that MMPI scales had little power for predicting adjustment to conditions of social isolation.

There were also no significant pre-post or

preisolation findings associated with the Lykken Sociopathy Scale or the Edwards Personal Preference Schedule.

Desocialization and Social Isolation

During the last experimental run, two dyads under similar experimental conditions completely deteriorated by the third day. In violation of mission procedures, these groups established verbal communication and jointly agreed to abort the mission. The disruptive behaviors accompanying this act occurred for a period of several hours. Their behavior included shouting obscenities through the ventilating system into other rooms, beating and kicking the outside door to their compartments, tampering with experimental equipment, refusing to execute assigned tasks properly, and generally ignoring instructions

from the mission control center. Because of the extreme nature of these violations, these groups were not included in the regular data analyses.

In order to investigate personality variables possibly related to disruption and maladaptation to social isolation, an equal number of so-called "riot-abort," abort, and completer groups under similar experimental conditions were selected for comparison. Table 4 presents premeans and postmeans and Kruskal-Wallis test (see Siegel, 1965, p. 184) results for these groups on a number of personality traits. There were seven traits on which these groups differed prior to confinement. The direction of these differences, confirmed by Mann-Whitney *U* tests, indicated that the riot-abort groups were significantly higher than completers and/or aborters on Repression-Sensitization (*R-S*), Control (*Cn*), Heterosexual, Aggression, Subjective Stress States (*SSS*), and Anger as measured by the Primary Affect Scale (Johnson & Myers, 1967). They were lower than completers and aborters on Responsibility (*Re*). Riot-aborters were also higher on several clinical scales, although differences were not statistically significant. Thus, the clinical instruments used in this study seemed to have good predictive ability in the assessment of maladjustment to social isolation, when the maladjustment is extreme.

Preisolation and postisolation analyses showed the riot-abort groups having elevated scores on a majority of the MMPI clinical scales whereas the completer and abort groups showed no such elevations. Riot-aborters had significantly higher postisolation scores on the MMPI Validity scale (*F*) ($p < .05$), Hypochondriasis ($p < .01$), Depression ($p < .10$), Paranoia ($p < .01$), Psychasthenia ($p < .01$), Schizophrenia ($p < .01$), and Repression-Sensitization ($p < .05$) scales. *T* scores on three of these scales (Hypochondriasis, Paranoia, and Schizophrenia) exceeded 70. Completer groups remained higher than the other two groups on Responsibility ($p < .10$). The Lykken Sociopathy Scale failed to distinguish between the groups prior to confinement; however, the means were in the expected direction. This trend continued following iso-

lation, with the riot-abort groups having a significantly greater mean on primary sociopathy ($p < .10$). Finally, the riot-abort and the abort groups were significantly higher than completer groups on postconfinement Subjective Stress States ($p < .10$), Anger ($p < .01$), and Animal ($p < .05$) on the HIT. Thus, these unusual groups not only had severe adjustment problems, but showed marked changes on most of the personality scales.

DISCUSSION

The present study was concerned with (a) the relationship between personality changes and various types of confinement and (b) the prediction of adjustment to isolation and confinement.

The major Holtzman Inkblot findings involved pre-post changes occurring for all groups, that is, decreases in *FD*, *M*, *I*, *H*, *Pn*, and Response Length and increases in *C*, *Sk*, and *L*. Similar findings for *FD*, *M*, *L*, and *C* were found by Cole et al. (1967) for both control and isolation groups, while a decrease in *H* occurred for isolated groups only. Studies reporting factor analyses of HIT variables (Gorham, 1967; Holtzman et al., 1961) found that *M*, *I*, *H*, and *FD* define a factor which Holtzman et al. (1961, p. 171) describe as dealing with ideational activity and imaginative capacity. That is, responses to inkblots in this cluster are felt to have little psychophysical correspondence in the stimuli but rather reflect the fantasy activity of the respondent. This factor accounted for the greatest proportion of the variance in our data. All of the variables in this factor showed postconfinement decreases, suggesting that social isolation produces an impoverishment of ideational and imaginative output. This interpretation is consistent with studies of sensory deprivation reporting postisolation breakdowns in thought processes, problem solving, and intelligence test performance (Bexton et al., 1954). Decreases in Human responses also confirm the Phillips and Smith (1953) hypothesis that the number of Human responses varies inversely with social isolation although Fernald and Linden (1966) failed to obtain such decrements.

A possible problem in interpreting the present data revolves around the question of pre-post HIT changes simply as a function of repeated testing rather than experimental conditions. Holtzman et al. (1961, p. 141), using a somewhat comparable sample, reported *slight* decreases in *Sh*, *M*, and *I*, simply as a function of repeated testing. The same authors found significant decreases on *M*, *I*, and *Pn*, and a significant increase in Location on two repeated testings (p. 133). However, the mean changes found in the present study were substantially larger than those of other studies and occurred on many more determinants. Furthermore, Cole et al. (1967) found no changes on these variables in an ambulatory control group. Finally, recent data from several studies conducted separately by Gorham and Holtzman on test-retest reliability, counterbalancing for form and length, indicate quite stable means for all variables.⁶ Thus, the evidence suggests that the magnitude and extent of HIT changes in the present study can be attributed to other than simple test-retest changes.

On a somewhat related issue, Megargee (1966) found that Response Length is highly loaded on the determinants in Factor I, and suggested that response length may be mediating changes in these variables. Support for this position can be seen in a study by Conners (1965) who found a highly significant increase in Factor I for an experimental (drug) group, while a placebo group showed no such change. One interpretation for change in this experimental group suggested that Response Length may have mediated this change. This notion was strengthened by the knowledge that one effect of the drug (dextroamphetamine) seems to be an increase in verbalization. In the present study a decrease in Factor I was accompanied by a decrease in total words used.

A second cluster contained two variables—Color and Shading, both of which reflect a sensitivity or concern with the stimulus qualities of the inkblots. Shading and Color responses increased from preisolation to post-isolation, especially for aborters. Increases in

Color have been interpreted as impulsiveness (Murray & Jackson, 1964). Cole et al. (1967) found increases in Color accompanied by decreases in Movement which they interpreted to mean that Ss were being more sensitive to external stimuli, while using fewer internal resources as the stress of isolation escalated. Data from the present study provide support for the finding and interpretation made by Cole et al. (1967). This interpretation is also strengthened in that similar changes occurred among the most stressful experimental conditions. That is, a significant increase in Color occurred for 20-day privacy groups, who also reported increasingly greater levels of subjective stress and anxiety over time and who had the highest abort rate of all experimental conditions (Taylor et al., 1968). The relationship between these changes and stress and anxiety results provide further support for a linkage between Color and Shading and certain affective states, such as anxiety and impulsiveness (Ainsworth, Klopfer, & Holt, 1954; Rapaport, Gill, & Schafer, 1946; Rorschach, 1949; Schachtel, 1943). However, these findings are not wholly in accord with the experimental literature which seems to indicate little relationship between Color on the Rorschach and anxiety (Cerbis & Nichols, 1963). The discrepancy might be due to the fact that the present study dealt with state anxiety as opposed to trait anxiety (see Taylor et al., 1968), while many studies in the literature have failed to make this distinction (Spielberger, 1966).

The attempt to find personality traits that would predict adjustment or success to social isolation met with only limited success. None of the instruments used in the present study detected preexperimental differences between aborters and completers. However, comparisons between extremely disruptive groups who aborted and groups confined under similarly stressful conditions revealed several personality traits which may be predictive of adjustment.

During the course of the experimental runs, two dyads became disruptive to the point of mutually aborting the mission. These disruptive groups differed from matched comparisons (completers and aborters) on several

⁶ Wayne H. Holtzman and Donald Gorham, personal communications, June, 1968.

traits. MMPI scores on Control and Responsibility for riot-aborters agreed with findings by Shipman (1963) indicating that tendencies to tempestuous and combative behavior appeared in people with low Responsibility and high Control scores. According to the EPPS, riot-aborters tended to be more aggressive and were significantly higher than other groups on expressions of heterosexual activities. The disruptive groups also had high Anger scores and were significantly lower than other groups on the HIT Barrier score. Megargee (1965) found that juvenile delinquents had Barrier scores significantly lower than nondelinquent adolescents. Furthermore, the more seriously delinquent Ss had Barrier scores significantly lower than the less seriously delinquent Ss. His interpretation and finding are consistent with events that led to the riotlike behavior described above.

A markedly noticeable difference between the disruptive groups and the controls was the extreme postconfinement elevation of the MMPI *F* (validity) scale, which seems to agree with other work on aggressive criminals (Gynther, 1961). However, the extremely elevated postisolation *F* score, which affects other MMPI scales and is usually interpreted as a "faked bad" record, leads us to view the MMPI changes as a deliberate effort by these groups to justify their disruptive behavior during confinement. Other changes following confinement indicated that the riot-abort groups tended to have elevated sociopathy scores and were more stressed and angered than other groups. Completers were sensitizers who admitted, prior to isolation, their concerns about being confined. Additionally, they were able to express aggression and anger and yet were not hostile.

Data reported elsewhere (see Taylor et al., 1968) showed that long missions with privacy and no stimulation from the outside constituted the most stressful conditions. A major finding in the present study was a significant change in perceptual responses from groups under these most stressful conditions. Specifically, the predicted changes in Color and Movement responses on the HIT occurred for groups under long mission expectation and privacy. This finding replicates a similar

change among incompatible dyads from the Cole et al. (1967) study, and relates to other findings in the literature suggesting a relationship between Color change and heightened affect. Pre-post changes reported here also replicated those reported by Cole et al. However, those changes associated with Factor I are confounded with decreased verbal productivity. Hence, decreases in *TWU* may have mediated decreases in Factor I scores.

The present study demonstrates that combinations of social isolation yielding varying degrees of stress, produce perceptual changes related to personality functioning. In addition to replicating an earlier finding relating increased stress to changes in Color and Movement, these data conform to a developing pattern of perception, personality, and overt behavioral response to the stressful effects of social isolation and confinement (Taylor et al., 1968). Post-hoc analyses of selected groups under the most extreme conditions of isolation revealed a syndrome of adjustment. Good adjustment was associated with flexibility, low sociopathy, and ability to be aggressive and angry yet not hostile. The completers were sensitizers who admitted freely to concern about being isolated. Further research is needed to define more definitely this adjustment syndrome. Results of such an effort will likely lead to accurate diagnosis of criteria related to adjustment in isolated situations, with implications for selection and training.

REFERENCES

- BEXTON, W. H., HERON, W., & SCOTT, T. H. Effects of decreased variation in sensory environment. *Canadian Journal of Psychology*, 1954, 8, 70-76.
- CERBUS, G., & NICHOLS, R. C. Personality variables and response to color. *Psychological Bulletin*, 1963, 60, 566-575.
- COLE, J. D., MACHIR, D., ALTMAN, I., HAYTHORN, W. W., & WAGNER, C. M. Perceptual changes in social isolation and confinement. *Journal of Clinical Psychology*, 1967, 23, 330-333.
- CONNERS, C. K. Effects of brief psychotherapy, drugs and type of disturbance on Holtzman inkblot scores in children. In *Proceedings of the 73rd annual convention of the American Psychological Association*. Washington, D. C.: American Psychological Association, 1965.
- EDWARDS, A. L. *Edwards Personal Preference Schedule manual*. (Rev. ed.) New York: Psychological Corporation, 1959.

- FERNALD, P. S., & LINDEN, J. D. The human content response in the Holtzman Inkblot Technique. *Journal of Projective Techniques and Personality Assessment*, 1966, 30, 441-446.
- FULKERSON, S. C., FREUD, S. L., & RAYNOR, G. H. The use of the MMPI in the psychological evaluation of pilots. *Journal of Aviation Medicine*, 1958, 29, 122-129.
- GORHAM, D. R. Validity and reliability studies of a computer-based scoring system for inkblot responses. *Journal of Consulting Psychology*, 1967, 31, 65-70.
- GYNTHIER, M. D. The clinical utility of "invalid" MMPI F scores. *Journal of Consulting Psychology*, 1961, 25, 540-542.
- HATHAWAY, S. R., & MCKINLEY, M. D. *Minnesota Multiphasic Personality Inventory*. (Rev. ed.) New York: Psychological Corporation, 1951.
- HAYTHORN, W. W., ALTMAN, I., & MYERS, T. I. Emotional symptomatology and subjective stress in isolated pairs of men. *Journal of Experimental Research in Personality*, 1966, 1, 290-305.
- HOLTZMAN, W. H., THORPE, J. S., SWARTZ, J. D., & HERRON, E. W. *Inkblot perception and personality*. Austin, Texas: University of Texas Press, 1961.
- HULL, J., & ZUBEK, J. P. Personality characteristics of successful and unsuccessful sensory isolation subjects. *Perceptual and Motor Skills*, 1962, 14, 231-240.
- JOHNSON, E., & MYERS, T. I. The development and use of the Primary Affect Scale (PAS). Technical Report No. 31, July 1967, Naval Medical Research Institute, Navy Contract MF022.01.03-1006.
- KLOFFER, B., AINSWORTH, M. D., KLOFFER, W., & HOLT, R. R. *Developments in the Rorschach technique*. Vol. 1. New York: Harcourt, Brace & World, 1954.
- LYKKEEN, D. T. A study of anxiety in the sociopathic personality. *Journal of Abnormal and Social Psychology*, 1957, 55, 6-10.
- MEGARGEE, E. I. Relation between barrier scores and aggressive behavior. *Journal of Abnormal Psychology*, 1965, 70, 307-311.
- MEGARGEE, E. I. The relation of response length to Holtzman Inkblot Technique. *Journal of Consulting Psychology*, 1966, 30, 415-419.
- MURRAY, J. E., & JACKSON, D. N. Impulsivity and color-form abstraction. *Journal of Consulting Psychology*, 1964, 28, 518-522.
- MYERS, T. I. Tolerance for sensory deprivation. In J. P. Zubek (Ed.), *Sensory deprivation: Fifteen years of research*. New York: Appleton-Century-Crofts, 1969.
- PHILLIPS, L., & SMITH, J. G. *Rorschach interpretation: Advanced technique*. New York: Grune & Stratton.
- RAPAPORT, D., GILL, M., & SCHAFER, B. S. *Diagnostic psychological testing*. Vol. 2. Chicago: Year Book Publishers, 1946.
- RORSCHACH, H. (Orig. publ. 1921) *Psychodiagnostics*. (Trans. by P. Lemkau & B. Kronenberg) New York: Grune & Stratton, 1949.
- SCHACTEL, E. On color and effect. *Psychiatry*, 1943, 6, 393-409.
- SELLS, S. B. (Ed.) "Tri-Service": Conference on research relevant to behavior problems of small military groups under isolation and stress. United States Air Force Arctic Aeromedical Laboratory, 1961, Air Force Contract 41 (657)323.
- SHIPMAN, W. G. The validity of MMPI hostility scales. Paper presented at the meeting of the American Psychological Association, Philadelphia, September 1963.
- SIEGEL, S. *Nonparametric statistics*. New York: McGraw-Hill, 1956.
- SPIELBERGER, C. D. Theory and research on anxiety. In C. D. Spielberger (Ed.), *Anxiety and behavior*. New York: Academic Press, 1966.
- TAYLOR, D. A., WHEELER, L., & ALTMAN, I. Stress reactions in socially isolated groups. *Journal of Personality and Social Psychology*, 1968, 9, 369-376.
- WRIGHT, M. W., CHYLINSKI, J., SISLER, G. C., & QUARRINGTON, B. Personality factors in the selection of civilians for isolated northern stations: A follow-up study. *Canadian Psychologist*, 1967, 8a, 23-31.
- WRIGHT, M. W., SISLER, G. C., & CHYLINSKI, J. Personality factors in selection of civilians for isolated northern stations. *Journal of Applied Psychology*, 1963, 47, 24-29.
- WRIGHT, N. A., & ZUBEK, J. P. Use of the multiple discriminant function in the prediction of perceptual deprivation tolerance. *Canadian Journal of Psychology*, 1966, 20, 105-113.

(Received July 12, 1968)

PSYCHOLOGICAL CHARACTERISTICS OF NEIGHBORHOOD YOUTH CORPS ENROLLEES¹

SHERMAN EISENTHAL² AND LEWIS J. SHERMAN³

Veterans Administration Hospital, Brockton, Massachusetts

The entire enrollment of an area Neighborhood Youth Corps (NYC) was administered the California Psychological Inventory (CPI) to facilitate program planning and evaluation. Compared to high school norms, all the enrollees showed significant achievement problems and social immaturity; the boys, however, were more extreme than the girls. They also showed a greater potential for dropping out of school and for acting-out delinquent behavior. The NYC results were also compared with norms developed on high school disciplinary problems and school dropouts.

The Neighborhood Youth Corps (NYC) is the work training program established by the Economic Opportunity Act as part of the war on poverty. A major objective of the program is to enhance the educational and vocational potential of its enrollees. These enrollees are of high school age and comprise a self-selected group from a disadvantaged economic background. The eligibility criteria, set by the Department of Labor, limit the maximum family income according to family size. In order to facilitate program planning and evaluation, it seems desirable to assess the psychological characteristics of this population.

Previous research on the lower social classes suggests an association with certain psychological variables: impulse expression, authoritarianism, low need achievement, high rate of school dropouts, etc. (Havighurst, 1949; Hollingshead, 1949; Leshan, 1952). Although enrollees are all "disadvantaged," they are also all self-selected. Hence it is not clear how representative they are of the low-income population from which they derive. There is at present no large body of data comparing

NYC enrollees and high school students along basic psychological dimensions aside from relatively nonrevealing demographic studies. This report attempts to describe the characteristics of the NYC enrollees with respect to two linked areas: (a) education and (b) social-emotional adjustment.

METHOD

The Ss were the entire enrollment in the Brockton, Massachusetts, NYC area during August 1967. There were 111 males and 150 females, ranging from the seventh to the twelfth grade, with most of the males in the ninth and tenth grades, and most of the females in the tenth and eleventh grades. They ranged in age from 14 to 21; most of the males were between 15 and 16 (55%) and most of the females were between 16 and 17 (63%). The enrollment was predominantly white (89%). The time in the program ranged from less than 1 month to over a year, the median for boys was 2 months and for girls was 3 months. Thirty-one of the enrollees, 12 male and 19 female, were high school dropouts. Since the distinction between "in school" and "out of school" is not a functionally sharp one, as a number of enrollees change their school status during the school year, the data from both groups were pooled for ease of presentation.

The enrollees were administered the California Psychological Inventory (CPI; Gough, 1964) and a 45-item questionnaire about the program. They were told that the purpose of the testing was for research directed toward better understanding of the enrollees and toward improving the effectiveness of the program. The CPI was selected because it is designed to measure many of the variables of special interest in this research, that is, achievement orientation, social adjustment, and intellectual efficiency. Furthermore, it is an instrument which has been widely used in counseling settings and has been standardized on populations possessing problems encountered in the NYC program, for example, adolescent delinquents and school dropouts. Finally,

¹ Another version of this paper was presented at the meeting of the Eastern Psychological Association, Washington, D. C., April 1968. The authors wish to thank Ray T. Mentzer, Executive Director of Self Help, Inc., and Vernon K. Sport, Administrator of the Brockton, Massachusetts, Neighborhood Youth Corps for their assistance.

² Requests for reprints should be sent to Sherman Eisenhal who is now at Boston University School of Medicine, 80 East Concord Street, Boston, Massachusetts 02118.

³ Now at the University of Missouri at St. Louis.

TABLE 1
COMPARISON OF NYC ENROLLEES AND HIGH SCHOOL STUDENTS ON CPI

Scale	Male					Female				
	NYC		High school		D_M	NYC		High school		D_M
	M	SD	M	SD		M	SD	M	SD	
Do	22.76	4.97	23.2	6.0	-.44	22.69	5.25	23.7	6.1	-1.01*
Cs	14.36	3.93	15.3	4.4	-.94*	13.95	3.79	16.0	4.9	-2.05**
Sy	20.65	4.01	21.5	5.4	-.85*	21.32	4.81	21.4	5.7	-.08
Sp	31.63	4.41	32.7	5.7	-1.07*	31.00	5.56	31.1	5.8	-.10
Sa	18.02	3.61	18.7	4.1	-.68	18.56	3.77	18.9	4.4	-.34
Wb	26.04	6.94	33.5	5.6	-7.46**	30.88	6.87	34.6	5.7	-3.72**
Re	23.06	5.00	26.7	5.7	-3.64**	26.92	5.32	30.0	5.2	-3.08**
So	29.30	5.74	36.3	6.0	-7.00**	35.42	6.55	39.4	5.6	-3.98**
Sc	22.36	7.26	25.3	8.0	-2.94**	26.40	7.10	27.6	8.5	-1.20*
To	15.21	5.22	17.8	5.3	-2.59**	17.03	5.55	18.7	5.5	-1.67**
Gi	14.84	5.66	15.1	6.2	-.26	15.48	5.54	15.7	6.2	-.22
Cm	19.12	5.40	25.2	2.8	-6.08**	24.10	3.62	26.1	1.9	-2.00**
Ac	20.36	4.85	22.3	5.3	-1.94**	22.60	5.13	24.1	5.3	-1.50**
Ai	14.34	4.45	14.6	4.1	-.26	14.90	4.34	15.5	4.2	-.60
Ie	29.20	6.32	33.6	6.3	-4.40**	32.23	6.19	34.4	6.5	-2.17**
Py	10.19	2.70	9.2	2.6	.99**	9.60	2.53	8.7	2.6	.90**
Fx	9.47	3.92	9.1	3.4	.37	8.73	3.62	8.9	3.2	-.17
Fe	16.74	3.28	15.4	3.6	1.34**	22.91	3.46	24.1	3.5	-1.19**

Note.—NYC enrollees were from Brockton, Massachusetts, during August 1967; for males, $N = 111$, for females, $N = 150$. High school standardization sample was reported by Gough (1964); for males, $N = 3,572$; for females, $N = 4,056$.

* $p < .05$.
** $p < .01$.

elaborate norms are available on a variety of relevant samples for comparison with the NYC groups.

RESULTS AND DISCUSSION

Table 1 compares NYC enrollees with the national norms for high school students published by Gough (1964). It is evident that the NYC males scored below the high school norms on 15 of 18 scales, 11 of which reached statistical significance. The male enrollees thus performed consistently and significantly below the high school norms. Three aspects of the comparative profiles are particularly salient. First, the extremely low scores on Well-Being and Communality, below a T value of 25, suggest a large number of care-less, atypical, and possible faked-bad protocols. Second, the NYC enrollees deviated from high school students most strongly on two of the four groups of CPI scales—social maturity and achievement.⁴ On the social

maturity scales, the NYC scored significantly below the high school norms on Responsibility, Socialization, Self-Control, Tolerance, and Communality. On the achievement scales, the NYC males scored significantly below the high school norms on Achievement via Conformance and Intellectual Efficiency. The Socialization scale could be included in this achievement group of CPI scales because it is a good "dropout" predictor (Gough, 1966). The NYC performance thus strongly suggests their potential for school problems and their difficulty in adjusting to conventional "middle-class" school environments.

To evaluate better the nature of their potential school problems, the NYC scores were compared with norms published by Gough (1964) on high school students with disciplinary problems (Table 2), most of whom finish school, and with norms on high

⁴ Gough divides CPI scales into four broad psychological categories: I. Measures of Poise, Ascendancy, and Self-Assurance; II. Measures of Socialization,

Maturity, and Responsibility; III. Measures of Achievement Potential and Intellectual Efficiency; IV. Measures of Intellectual and Interest Modes.

school dropouts (Gough, 1966) (Table 3). The enrollees appeared considerably more similar to the disciplinary problem group than to the high school students. The NYC pattern of scoring consistently below the high school norms on social maturity and achievement scales thus does not apply to the disciplinary problem group. Table 2 shows that the NYC enrollees scored below the disciplinary problem group on 10 of the 18 scales, 5 of which reached statistical significance. On the achievement scales, the NYC scored significantly higher on Achievement via Conformance, but significantly lower on Intellectual Efficiency. On the social maturity scales, the enrollees scored significantly lower on Socialization and Communality. This pattern suggests that the students with disciplinary problems have focused their rebellion in settings where achievement requires conformity, but are somewhat better organized and relatively more socially mature than the male NYC enrollees.

The scores of the NYC male enrollees appeared more similar to the dropout norms than to the high school norms. The enrollees scored below the dropouts on 10 of the 18 CPI scales, 5 of the 10 attaining statistical significance (Table 3). The NYC enrollees seemed to share the same difficulties with the dropouts in self-direction and achievement in a structured situation. The NYC enrollees, however, retained achievement values in unstructured situations to a greater degree than dropouts. Despite this, the enrollees tended to be more socially immature than dropouts. The fact that the male NYC enrollees scored significantly lower on the Socialization scale than the dropouts, in addition to the Responsibility and Self-Control scales, suggests that for a number of NYC enrollees their social immaturity may manifest itself in a more extreme manner than dropping out of school, perhaps in more unconventional, delinquent, and acting-out behavior.

TABLE 2
COMPARISON OF NYC ENROLLEES AND HIGH SCHOOL DISCIPLINARY
PROBLEM STUDENTS ON THE CPI

Scale	Male					Female				
	NYC		High school disciplinary problem		D_M	NYC		High school disciplinary problem		D_M
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Do	22.76	4.97	22.2	5.8	.56	22.69	5.25	22.8	5.3	-.11
Cs	14.36	3.93	15.0	4.2	-.64	13.95	3.79	15.4	4.6	-1.45*
Sy	20.65	4.01	21.4	5.1	-.75	21.32	4.81	20.7	4.5	.62
Sp	31.63	4.41	34.0	5.6	-2.37**	31.00	5.56	31.8	5.3	-.80
Sa	18.02	3.61	18.9	3.7	-.88	18.56	3.77	19.2	4.4	-.64
Wb	26.04	6.94	31.2	5.8	-5.16**	30.88	6.87	31.6	5.2	-.72
Re	23.06	5.00	21.7	4.7	1.36	26.92	5.32	26.0	5.6	.92
So	29.30	5.74	31.3	5.4	-2.00*	35.42	6.55	34.8	7.1	.62
Sc	22.36	7.26	21.7	6.9	.66	26.40	7.10	23.9	8.4	2.50*
To	15.21	5.22	15.8	4.8	-.59	17.03	5.55	16.2	5.1	.83
Gi	14.84	5.66	13.1	5.6	1.74*	15.48	5.54	14.0	5.4	1.48*
Cm	19.12	5.40	23.8	3.2	-4.68**	24.10	3.62	25.5	2.1	-1.40**
Ac	20.36	4.85	18.7	4.3	1.66**	22.60	5.13	20.8	5.3	1.80*
Ai	14.34	4.45	13.2	3.9	1.14	14.90	4.34	14.3	4.0	.60
Ie	29.20	6.32	31.0	6.2	-1.80*	32.23	6.19	32.4	6.4	-.17
Py	10.19	2.70	8.7	2.3	1.49**	9.60	2.53	8.4	2.4	1.20**
Fx	9.47	3.92	9.7	2.7	-.23	8.73	3.62	9.0	3.2	-.27
Fe	16.74	3.28	13.7	3.3	3.04**	22.91	3.46	23.5	3.5	-.59*

Note.—NYC enrollees were from Brockton, Massachusetts, during August 1967; for males, $N = 111$; for females, $N = 150$. High School Disciplinary standardization sample was reported by Gough (1957); for males, $N = 91$; for females, $N = 87$.

* $p < .05$.

** $p < .01$.

To determine the degree of similarity among the groups, score differences were calculated between enrollee and high school student norms, enrollee and disciplinary problem norms, and enrollee and dropout norms, and the significance of these differences tested by the Wilcoxon matched-pairs signed-ranks test. The difference scores between the enrollees and the high school students differed significantly from those with the high school dropouts but not from those with the high school disciplinary problems. This finding serves to emphasize the relative similarity between the NYC male enrollees and high school dropouts.

The female NYC enrollees differed from the male enrollees more than they resembled them. Like the males, the females scored below the high school norms (Gough, 1964) (Table 1) on 17 of 18 scales, 11 of the 17 comparisons attaining statistical significance. Also, the female enrollees scored significantly below the norms in scales measuring social

maturity and achievement orientation. However, unlike the boys, the females did not obtain the extremely low scale scores. The females appeared much less deviant and disturbed than the males for most psychological dimensions.

The likelihood of the enrollees developing school problems was explored by comparing their scores with norms for disciplinary (Gough, 1964) (Table 2) and dropout groups (Gough, 1966) (Table 3). The results from the female enrollees resembled those of the disciplinary problem students more closely than those of the high school students. Furthermore, the female enrollees did not score consistently above or below the disciplinary students. Thus, the enrollees scored below the disciplinary group on only 9 of the 18 scales with just 2 of the 9 lower scores reaching statistical significance. Four of the 9 scales on which the enrollees scored above the disciplinary group reached statistical significance. A comparison of the two profile pat-

TABLE 3
COMPARISON OF NYC IN SCHOOL AND HIGH SCHOOL DROPOUTS ON THE CPI

Scale	Male					Female				
	NYC		High school dropouts		D_M	NYC		High school dropouts		D_M
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Do	22.64	5.12	22.54	5.77	.10	22.44	5.37	24.10	6.12	-1.66*
Cs	14.20	3.94	13.58	3.77	.60	14.21	3.80	14.99	4.43	-.78
Sy	20.58	4.00	20.84	5.24	-.26	21.38	4.89	20.91	5.62	.47
Sp	31.33	4.77	31.10	5.53	.23	30.59	5.46	30.38	5.98	.21
Sa	18.51	4.24	18.02	3.54	.49	18.51	3.76	18.42	4.24	.09
Wb	26.30	6.12	32.09	5.59	-5.79**	30.71	6.68	31.64	6.03	-.93
Re	23.23	4.77	24.73	4.90	-1.50*	27.20	5.36	28.86	4.91	-1.66**
So	29.41	5.77	34.62	5.61	-5.21**	35.27	6.31	37.61	5.89	-2.34**
Sc	21.90	7.10	24.28	8.18	-2.38*	26.28	7.28	25.88	8.04	.40
To	14.95	5.16	16.28	5.45	-1.33	17.24	5.68	16.55	5.07	.69
Gi	14.55	5.67	15.53	6.42	-.98	15.41	5.54	15.73	6.30	-.32
Cm	19.81	4.42	24.79	2.56	-4.98**	23.86	3.62	25.56	2.58	-1.80**
Ac	20.26	4.94	20.85	4.80	-.59	22.68	5.10	22.79	4.98	-.09
Ai	14.32	4.45	12.41	3.82	1.91**	14.99	4.35	14.00	3.89	.99*
Ie	29.40	6.53	30.86	6.15	-1.46	31.89	6.99	32.15	5.94	-.26
Py	10.03	2.54	8.80	2.54	1.23**	9.82	2.52	8.48	2.49	1.34**
Fx	9.43	3.79	8.34	3.23	1.09*	8.90	3.47	8.48	3.09	.42
Fe	16.73	3.15	16.08	3.45	.65	22.92	3.39	24.44	3.67	-1.52**

Note.—NYC enrollees were from Brockton, Massachusetts, during August 1967; for males, $N = 99$; for females, $N = 131$. High school dropout sample was reported by Gough (1966); for males, $N = 124$; for females, $N = 160$.

* $p < .05$.

** $p < .01$.

terns suggests (a) the NYC females are somewhat more socially mature though less confident of their social status than the disciplinary students and (b) the NYC females are relatively less conflicted and rebellious in school settings.

As was true of the comparison with the disciplinary students, the scores of female enrollees were closer to those of the dropouts than to the general high school student. The enrollees scored below the dropouts on 10 of the 18 CPI scales with 5 of the 10 reaching statistical significance (Table 3). On the social maturity scales, the enrollees scored significantly below the dropouts on Responsibility and Socialization. On the achievement scales, the enrollees scored significantly above the dropouts in Achievement via Independence but not Achievement via Conformance or Intellectual Efficiency. This is the same pattern found in the comparison of the NYC males and the dropout norms. These findings suggest that the female enrollees are relatively less mature than the female dropouts but place more intrinsic value on achievement than do the dropouts.

It is of interest to compare the NYC enrollees and the three other groups in terms of the CPI factor analysis reported by Nichols and Schnell (1963). They obtained two primary factors in their study which they labeled Value Orientation (Factor I) and Person Orientation (Factor II). Factor I measures "responsibility, self-control, and maturity as well as a tendency to minimize complaints and a concern over the reaction of others [pp. 230-231]." The enrollees differed from high school students primarily on this Factor I and consistently demonstrated greater social immaturity than the other groups.

In conclusion, then, there is an ironical contrast with regard to the sexes in terms of participation in the NYC. The boys, who show the greater need for help, are underrepre-

sented as compared with the girls. Attracting and keeping the male enrollees, who most need direction and assistance, remains a challenging but unresolved problem for program planners. The NYC boys show significant achievement problems and a high potential for dropping out of school. They are also socially immature and have considerable potential for unconventional and delinquent behavior. Their sense of intellectual inadequacy is striking. Their style is to act out their problems and behave in an impulsive, nonreflective and intellectually rigid pattern. The NYC girls also manifest a potential for achievement problems but would seem to be less likely to drop out of school than the boys. In addition, the girls are also more mature and less likely to engage in unconventional and delinquent behavior. The pattern of scores on achievement related scales suggests that the male NYC enrollees do have achievement values but that they lack the social maturity and intellectual efficiency to effectively perform in the school setting. The essential issue is how to cope with the male enrollee's impatience and lack of direction and recognize that he is likely to rebel in response to academic social pressure.

REFERENCES

- GOUGH, H. G. *Manual for the California Psychological Inventory* (Rev. ed.) Palo Alto, Calif.: Consulting Psychologists Press, 1964.
- GOUGH, H. G. Graduation from high school as predicted from the California Psychological Inventory. *Psychology in the Schools*, 1966, 3, 208-216.
- HAVIGHURST, R. J., & TABA, H. *Adolescent character and personality*. New York: Wiley, 1949.
- HOLLINGSHEAD, A. B. *Elmstown's youth*. New York: Wiley, 1949.
- LESHAN, L. L. Time orientation and social class. *Journal of Abnormal and Social Psychology*, 1952, 48, 589-592.
- NICHOLS, R. C., & SCHNELL, R. R. Factor scales for the California Psychological Inventory. *Journal of Consulting Psychology*, 1963, 27, 228-235.

(Received July 18, 1968)

EFFECTS OF PRESESSION INFORMATION ON PERCEPTION OF THE THERAPIST AND RECEPTIVITY TO INFLUENCE IN A PSYCHOTHERAPY ANALOGUE¹

ROGER P. GREENBERG²

*Department of Psychiatry, State University of New York
Medical School, Syracuse*

In order to investigate the idea that pre-session information can change Ss' evaluations and perceptions of a therapy session, 112 Ss were randomly assigned to one of four groups and given different information with regard to a therapist's warmth and experience. Following structuring, all Ss listened to the same tape of a therapy session. Measures obtained following the tape revealed that Ss structured warm or experienced, as opposed to cold or inexperienced, were more attracted to the therapist, more receptive to his influence, and evaluated his work more positively. Further, Ss structured warm, rather than cold, were more willing to meet with the therapist and more persuaded by his communications. The findings thus support the notion that premeeting information can affect the therapy relationship.

As the study of psychotherapy proceeds, more time and effort are being turned toward an examination of the therapy relationship in an attempt to clarify, understand, and eventually control those variables which would promote the practice of more efficient therapy. The psychotherapeutic relationship was the object of study in the present investigation. More specifically, interest was focused on the class of pretherapy or early therapy events which might advantageously affect therapeutic process and outcome. The present analogue study endeavored to show that information given prior to an S's exposure to a therapy session can change his perception of that session and render him more open and receptive to the therapist's influence attempts.

The fact that "structuring" can influence an S's perception of a person has been amply demonstrated by the impression-formation literature. A basic finding of this area of study seems to be that we react to a person as a unity or whole and not as a series of specific traits (Allport, 1937; Asch, 1946; Gollin,

1954). Further, what a person perceives will be determined by his total impression of what he feels he should see. That is, he will distort an "objective" encounter along lines that agree with his prior expectations and theories (e.g., Asch, 1946; Farina & Ring, 1965; Kelley, 1950; Secord, 1958). In an analogous fashion, the therapy literature seems to show that pretherapy expectancies or manipulations can affect the therapy relationship (e.g., Frank, 1959; Goldstein, 1962). Thus, since structuring or expectations do play an important role in interactions between people, it seems reasonable to ask which variables may act to enhance or diminish the quality of the initial therapy relationship.

Theorizing and common sense support the notion that therapist-patient interaction can proceed more efficiently in an atmosphere of cooperation than in one of hostility and resistance. Further, a number of studies have shown that a person is more susceptible to influence attempts made by an individual he is attracted to or likes (Back, 1951; Burdick & Burnes, 1958; Sapolsky, 1960). Therefore, the creation of an agreeable therapy relationship may be viewed as intimately related to variables which enhance patient attraction and foster receptivity to influence.

Two of the variables which have received considerable attention in discussions of ther-

¹ This article is based on a doctoral dissertation submitted in partial fulfillment of the requirements for the PhD degree at Syracuse University. The author wishes to express his gratitude to Arnold P. Goldstein for his aid and encouragement.

² Requests for reprints should be sent to Roger P. Greenberg, Department of Psychiatry, State University of New York at Syracuse, 750 East Adams Street, Syracuse, New York 13210.

apy are therapist "warmth" and "experience." Both of these variables have been hypothesized to be conducive to successful therapy outcome. Truax and Carkhuff (1967) have suggested in a host of studies that warmth is an essential element of successful therapy and quote from every major psychotherapy theorist in presenting their argument. The fact that individuals perceived as behaving in a "warm" manner can affect the behavior and attitudes of others has been experimentally demonstrated many times (e.g., Heller, Meyers, & Kline, 1963; Kelley, 1950; Masling, 1957). In like manner, it has been reported that the experience of the therapist may be an important indicator of potential therapeutic gain or adequacy (e.g., Barrett-Lennard, 1962; Cartwright & Vogel, 1960; Strupp, 1962).

The present study attempted to find out if just labeling a therapist "warm" or "experienced" can change the nature of the therapy relationship. Do Ss perceive a therapy session differently if they feel they are listening to a warm or experienced therapist? Does a therapist have an advantage in influencing patients merely by virtue of being labeled warm or experienced? Specifically, it was hypothesized that Ss structured "warm" would be more attracted to the therapist and more receptive to his influence attempts than Ss structured "cold." Similarly, it was hypothesized that Ss structured "experienced" would be more attracted to the therapist and receptive to his influence attempts than those structured "inexperienced." Above all, it was expected that Ss structured both "warm" and "experienced" would be more attracted to the therapist and receptive to his influence attempts than Ss structured both "cold" and "inexperienced."

METHOD

Subjects

The Ss were 112 undergraduates obtained from introductory psychology courses at Syracuse University. Each S was randomly assigned to one of four equal sized groups: experienced-warm, experienced-cold, inexperienced-warm, or inexperienced-cold. Each group contained 14 men and 14 women. The Ss were run in groups of 12-16 with each condition being run on two separate occasions.

Materials

After initial structuring all Ss listened to a 15-minute tape of a simulated therapy session. This tape was judged by 10 advanced clinical psychology graduate students and three PhD clinical psychologists. All agreed that the tape was a credible representation of a psychotherapy session and relatively ambiguous with regard to the variables of "warmth" and "experience."

Following the tape presentation Ss were asked to fill out two questionnaires. Questionnaire 1 consisted of a scale to measure each S's attraction toward the taped therapist and a scale to measure each S's receptivity to the taped therapist's influence. The attraction scale was a modification of the Client's Personal Reaction Questionnaire developed by Ashby, Ford, Guernsey, and Guernsey (1957), and meaningfully used in studies of the therapy relationship by Goldstein (1962) and Snyder (1961). The scale consists of a series of 28 statements describing positive and negative aspects of the taped therapist-patient interactions, to which each S responded on a 7-point scale from "strongly agree" to "strongly disagree." A constant of 84 was added to each S's score so that the possible range of attraction scores was 0-168. Ten statements concerning Ss' own attitudes toward the therapist made up the receptivity scale (e.g., I would probably feel uncomfortable talking with this therapist). Again Ss rated their degree of agreement with each statement on a 7-point scale. Scores were taken to reflect each S's willingness to interact and be open with the taped therapist. A constant of 30 was added to each S's score so that the possible range of receptivity to influence scores was 0-60.

Questionnaire 2 was a persuasibility measure and attempted to measure the therapist's ability to influence Ss' perceptions and opinions of the taped session. The Ss were presented with a series of 40 "psychologically based" statements about the taped patient (e.g., the patient is probably absentminded) and asked to rate these statements on a 7-point scale from "strongly agree" to "strongly disagree." All answers were matters of opinion since there was no factual information in the tape on which to base answers to these questions. The Ss were also instructed that the taped therapist had filled out this questionnaire and were given his rating after each item so that they could "compare" their views with his. The therapist ratings presented were randomly distributed (i.e., slightly agree through slightly disagree) except for 20 items which were randomly rated at the extremes (i.e., strongly agree or strongly disagree). The Ss were scored for each of the 20 selected extreme items on a scale ranging from 1-7, with 7 representing perfect correspondence with the "therapist's opinion"; thus the possible range of persuasibility scores was 20-140. The higher the score for any S the more his ratings coincided with those of the therapist and thus, presumably the more his views were influenced by the therapist.

Finally, Ss were presented with a printed sheet stating that the psychology department, after many inquiries concerning student problems, was thinking of asking the taped therapist to the university to discuss problems. The Ss were then asked in writing if they were willing or not willing to meet the taped therapist to discuss problems.

Two questions in Questionnaire 1 served as a check on the independent warmth and experience manipulations (i.e., the therapist is a warm person; the therapist sounded as though he had dealt with the problems of many people in the past).

Procedure

After each group of Ss was seated in the experimental room, they were presented with a printed page headed "Biographical Information." The E then read the printed page to the Ss. The information given was as follows:

In recent years more and more people experiencing emotional and psychological problems have been seeking help from psychotherapists. We are interested in learning more about psychotherapy and the reactions of people to this treatment. Therefore, we are going to play a tape of part of an actual psychotherapy session that took place in a large city other than Syracuse and ask you for your reactions to it. As you listen to the tape we would like you to put yourself in the place of the patient and imagine how you, as the patient, would react to this particular session.

The patient in this session is a college student at a large university who sought help after finding, among other things, that he was having increasing difficulty making decisions.

The rest of the information given to each S varied depending on which of the four groups he had been assigned to.

Experienced: warm or cold. The therapist has been engaged in the practice of therapy for over 20 years and has lectured and taught at some of the country's leading universities and medical schools. Questionnaires submitted to the therapist's colleagues seem to reveal that he is a rather (warm or cold) person, industrious, critical, practical, and determined.

Remember to try and put yourself in the patient's place as you listen to the tape.

Inexperienced: warm or cold. The therapist is a student engaged in a required part of his training and this session is one of his first experiences as a therapist. Questionnaires submitted to the therapist's colleagues seem to reveal that he is a rather (warm or cold) person, industrious, critical, practical, and determined.

Remember to try and put yourself in the patient's place as you listen to the tape.

Following the playing of the tape, all Ss were asked to fill out the questionnaires previously described.

RESULTS

Before the main hypotheses were tested, the data for the questions serving as a check on the independent warmth and experience manipulations were examined by two-tailed t tests for mean differences. Analysis revealed that Ss structured warm rated the taped therapist as significantly warmer than Ss structured cold ($t = 10.51$, $df = 110$, $p < .001$). Similarly, Ss structured experienced, as opposed to inexperienced, rated the taped therapist as sounding significantly more experienced ($t = 5.98$, $df = 110$, $p < .001$). Thus, both the warmth and experience manipulations were successful.

The mean attraction, receptivity, and persuasibility scores are presented in Table 1. Separate analyses of variance applied to these three measures showed, as predicted, that Ss structured warm, rather than cold, were more attracted to the therapist ($F = 50.68$, $df = 1/104$, $p < .001$), more receptive to his influence ($F = 28.15$, $df = 1/104$, $p < .001$), and more persuaded by his communications ($F = 7.06$, $df = 1/104$, $p < .01$). Also as predicted, Ss structured experienced were more attracted ($F = 8.00$, $df = 1/104$, $p < .01$) and receptive ($F = 8.21$, $df = 1/104$, $p < .01$) to the therapist than Ss structured inexperienced. Yet, Ss structured experienced were not more easily persuaded by the therapist's statements. Scheffé's tests for mean differences applied to the three measures revealed, as predicted, the warm-experienced group was more attracted ($p < .001$), receptive ($p < .001$), and persuaded ($p < .05$) by the taped therapist than the cold-inexperienced group. In addition, Ss given the information that the therapist was warm, rather than cold, were more attracted to him regardless of

TABLE 1
MEAN ATTRACTION, RECEPTIVITY, AND
PERSUASIBILITY SCORES

Information given	Attraction		Receptivity		Persuasibility	
	Warm	Cold	Warm	Cold	Warm	Cold
Experienced	118.25	83.18	39.46	27.21	103.54	99.43
Inexperienced	105.43	66.68	32.86	20.57	101.96	93.86

whether the "warm" and "cold" were paired with "experience" or "inexperience" (p 's < .001). The Ss were also more receptive to the warm therapist regardless of whether the "warm" and "cold" were paired with "experience" or "inexperience" (p 's < .01). Finally, Ss were more attracted to the therapist presented as warm-inexperienced rather than cold-experienced (p < .05).

Three chi-square analyses were done on the gross measure of whether or not each S was willing to meet with the taped therapist to discuss student problems. As predicted, a significantly larger number of Ss structured warm (42), as compared to cold (30), stated they would be willing to meet with the taped therapist ($\chi^2 = 4.71$, $df = 1$, $p < .05$). Also, as predicted, a significantly larger number of Ss structured warm-experienced (23), as opposed to cold-inexperienced (14), stated they would be willing to meet with the taped therapist ($\chi^2 = 5.10$, $df = 1$, $p < .05$). No significant difference in willingness to meet with the therapist was obtained between Ss structured experienced and those structured inexperienced.

DISCUSSION

In general then, both the trait (i.e., warm-cold) and the role (i.e., experience-inexperience) information presented seemed to affect Ss' perceptions and evaluations of the session. However, the trait information was apparently more powerful in biasing perceptions, in this sample, than the role information. In evaluating the results, it should be kept in mind that the present study is a psychotherapy analogue and as such did not necessarily use individuals with psychological problems as Ss. Therefore, it seems tenable that the experience dimension may not have been as important a dimension for this group as for a group of actual patients seeking help. The warmth dimension, on the other hand, seems important as a facilitator of rapport for people involved in any kind of relationship. Also, since both the experienced and novice therapist were presented to Ss as trained individuals the effects of the "experience" information may have been attenuated within the present study. In like manner, these effects may be

attenuated within the practice of psychotherapy at large where patients are seen by therapists known to have varying degrees of experience. Further, since most people do not know how a therapist behaves or what goes on in a psychotherapy session, it may be more difficult for Ss to bias their perceptions given the information that a therapist is experienced than given the information that he is warm. The impression-formation literature makes the point that negative information about a person has more influence on Ss' perceptions of that person than positive information (e.g., Briscoe, Woodyard, & Shaw, 1967; Pastore, 1960). Here it is probable that Ss would feel more resentful and negative toward a cold therapist than an inexperienced one since it is assumed that individuals have control over acting warm or cold but not over being experienced or inexperienced.

Finally, and perhaps most significant, the traits of the therapist (i.e., warmth) may be much more important than the number of hours of his experience or his role title in promoting therapeutic change. This idea seems to be gaining support as studies point out that lay people, with certain attributes, can be trained in a relatively short period of time to do effective therapy (e.g., Carkhuff & Truax, 1965; Rioch, Elkes, Flint, Udansky, Newman, & Silber, 1963).

The present study suggests that premeeting information may significantly affect the process of psychotherapy. It highlights the impact which referral sources, rumor, or preconceived notions may have on psychotherapy. Further, it underscores the importance of the patient's perception of the therapist for facilitation of attraction and influence. It is planned that future research will move further along the analogical bridge toward the actual therapy room with actual therapy patients. The findings of the present study strongly suggest that "structuring" can enhance the initial psychotherapeutic relationship.

REFERENCES

- ALLPORT, G. W. *Personality*. Boston: Houghton Mifflin, 1937.
- ASCH, S. E. Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 1946, 41, 258-290.

- ASHBY, J. D., FORD, D. H., GUERNEY, B. G., & GUERNEY, L. F. Effects on clients of a reflective and a leading type of psychotherapy. *Psychological Monographs*, 1957, 71(24, Whole No. 453).
- BACK, K. W. Influence through social communication. *Journal of Abnormal and Social Psychology*, 1951, 46, 9-23.
- BARRETT-LENNARD, G. T. Dimensions of the client's experience of his therapist associated with personality change. *Genetic Psychology Monographs*, 1962, 76, No. 43.
- BRISCOE, M. E., WOODYARD, H. D., & SHAW, M. E. Personality impression change as a function of the favorableness of first impressions. *Journal of Personality*, 1967, 35, 343-357.
- BURDICK, H. A., & BURNES, A. J. A test of "strain toward symmetry" theories. *Journal of Abnormal and Social Psychology*, 1958, 57, 367-369.
- CARKEHUFF, R. R., & TRUAX, C. B. Training in counseling and psychotherapy: An evaluation of an integrated didactic and experiential approach. *Journal of Consulting Psychology*, 1965, 29, 333-336.
- CARTWRIGHT, R. D., & VOGEL, J. L. A. A comparison of changes in psychoneurotic patients during matched periods of therapy and no therapy. *Journal of Consulting Psychology*, 1960, 24, 121-127.
- FARINA, A., & RING, K. The influence of perceived mental illness on interpersonal relations. *Journal of Abnormal Psychology*, 1965, 70, 47-51.
- FRANK, J. D. The dynamics of the psychotherapeutic relationship. *Psychiatry*, 1959, 22, 17-39.
- GOLDSTEIN, A. P. *Therapist-patient expectancies in psychotherapy*. New York: Pergamon Press, 1962.
- GOLLIN, E. Forming impressions of personality. *Journal of Personality*, 1954, 23, 65-76.
- HELLER, K., MYERS, R. A., & KLINE, L. V. Inter-viewer behavior as a function of standardized client roles. *Journal of Consulting Psychology*, 1963, 27, 117-122.
- KELLEY, H. H. Warm-cold variable in first impressions. *Journal of Personality*, 1950, 18, 431-439.
- MASLING, J. M. The effects of warm and cold interaction on the interpretation of a projective protocol. *Journal of Projective Techniques*, 1957, 21, 377-383.
- PASTORE, N. Attributed characteristics of liked and disliked persons. *Journal of Social Psychology*, 1960, 52, 157-163.
- RIOCH, M. J., ELKES, C., FLINT, A. A., UDANSKY, B. S., NEWMAN, R. G., & SILBER, E. National Institute of Mental Health pilot study in training mental health counselors. *American Journal of Orthopsychiatry*, 1963, 33, 678-689.
- SAPOLSKY, A. Effect of interpersonal relationships upon verbal conditioning. *Journal of Abnormal and Social Psychology*, 1960, 60, 241-246.
- SECORD, P. F. Facial features and inference processes in interpersonal perception. In R. Taguiri & L. Petrillo (Eds.), *Person perception and interpersonal behavior*. Stanford: Stanford University Press, 1958.
- SNYDER, W. U. *The psychotherapy relationship*. New York: Macmillan, 1961.
- STRUPP, H. H. The therapists' contribution to the treatment process. In H. H. Strupp & L. Luborsky (Eds.), *Research in psychotherapy*. Vol. 2. Washington, D. C.: American Psychological Association, 1962.
- TRUAX, C. B., & CARKEHUFF, R. B. *Toward effective counseling and psychotherapy: Training and practice*. Chicago: Aldine, 1967.

(Received July 19, 1968)

DIGIT SPAN:

AN INDICANT OF TRAIT OR STATE ANXIETY?¹

WILLIAM F. HODGES²

AND

CHARLES D. SPIELBERGER

University of Colorado

Florida State University

Clinical psychologists generally interpret lower scores on Digit Span (DS) relative to other Wechsler subtests as indicating anxiety. However, research findings on this topic are inconsistent and difficult to interpret because of ambiguity with regard to the concept of anxiety. To clarify this ambiguity, Spielberger has proposed a conception of anxiety that elaborates on the distinction, initially proposed by Cattell and Scheier, between anxiety as a transitory state (A-state) and as a relatively stable personality trait (A-trait). The present study evaluated the relationship between DS performance and measures of A-trait (Taylor Manifest Anxiety scale) and A-state (Zuckerman Adjective Checklist) for 72 male undergraduates. The Ss reporting high levels of A-state showed significant decrements in DS performance. There was no difference in the DS performance of high- and low-A-trait Ss.

For many years clinical psychologists have interpreted the pattern of scores obtained on subtests of the Wechsler Intelligence Scale as reflecting important personality characteristics of individuals taking the test. One of the most frequent inferences is that a low score on the Digit Span (DS) test, as contrasted with scores on other subscales, indicates the presence of anxiety (Rapaport, Gill, & Schafer, 1968; Rapaport, Schafer, & Gill, 1944; Wechsler, 1958).

While clinicians tend to agree that DS performance is lowered by "anxiety," the meaning and conceptual status of anxiety is ambiguous. Cattell and Scheier (1961) have noted that the term "anxiety" has been used to refer both to a transitory state or condition of the organism and to a relatively stable personality trait. Although variables that loaded differentially on their state and trait anxiety factors have been identified by Cattell and Scheier, they have not specified the nature of

the relationship between these two anxiety concepts.

Spielberger has recently proposed a theoretical conception of anxiety that specifies the relationship between the concepts of state and trait anxiety (Spielberger, 1966; Spielberger & Lushene, in press). Building on the work of Freud (1933, 1936), Cattell and Scheier (1961), and Spence (1964), Spielberger defines state anxiety (A-state) as consisting of subjective feelings of apprehension and concern and heightened autonomic nervous system activity; whereas trait anxiety (A-trait) refers to anxiety proneness, that is, to individual differences in the disposition to respond with high levels of A-state under stressful circumstances.

The failure in most investigations of DS and anxiety to distinguish between A-state and A-trait has made the literature in this area difficult to interpret. Anxiety (presumably A-state) was induced in two studies by telling Ss they were doing poorly on a test of ability (Griffiths, 1958; Moldawsky & Moldawsky, 1952). For Ss exposed to this stressful experimental condition, significant decrements in DS performance were obtained as compared to nonstress control groups. In two other studies in which the anxiety measures could be reasonably interpreted as indicants of A-state, decrements in DS performance were also found to be associated with increases

¹ The research summarized in this report was based on data obtained while the authors were associated with the Department of Psychology at Vanderbilt University. It was supported in part by a United States Public Health Service Fellowship (5-F1-MH-28,537) from the National Institute of Mental Health to the first author and by a research grant (HD 947) to the second author from the National Institute of Child Health and Human Development, United States Public Health Service.

² Requests for reprints should be sent to William F. Hodges, Department of Psychology, University of Colorado, Boulder, Colorado 80302.

in A-state (Pyke & Agnew, 1963; Walker & Spence, 1964).

The results have been inconsistent and contradictory in investigations of the relationship between DS and presumed measures of A-trait, such as the Taylor Manifest Anxiety (MA) scale, the IPAT Anxiety scale, and the Welsh A scale. Three studies have reported no relationship between DS performance and A-trait measures (Jackson & Bloomberg, 1958; Matarazzo, 1955; Walker & Spence, 1964), one study found a positive relationship (Jurjevich, 1963), and one reported a negative relationship (Calvin, Koons, Bingham, & Fink, 1955). The results have also been equivocal in studies in which clinical diagnosis was used as the criterion of A-trait (Lewinski, 1945; Merrill & Heathers, 1952; Warner, 1950).

The present study was designed to evaluate the relationship between DS performance and measures of state and trait anxiety. The Ss who differed in A-trait were required to perform on a DS task under failure and control conditions.⁸ In accordance with Trait-State Anxiety Theory (Spielberger & Lushene, in press), high-A-trait Ss were expected to respond with higher levels of A-state than low-A-trait Ss in the failure condition, but not in the nonthreatening control condition. It was further expected that performance decrements in DS would be inversely related to level of A-state but not necessarily to A-trait, since anxiety must be aroused in order for it to interfere with behavior.

METHOD

The Ss were 72 male undergraduates enrolled in the introductory psychology course at Vanderbilt University who volunteered for an experiment on "numbers." The Taylor (1953) MA scale served as the measure of A-trait. The Ss with MA scale scores of 18 or greater were designated the high-A-trait group; those with scores of 11 or lower were designated the low-A-trait group. These cutoff points approximated the upper and lower quartiles of the MA scale distribution for males, after Ss with Lie scale scores of 8 or higher had been eliminated from the S pool.

⁸ An additional condition involving the threat of shock was also used, but these data were not relevant to the hypotheses of the present study and are reported elsewhere (Hodges, 1968).

Prior to the experiment, each S was assigned either to the Failure (stress) Condition or the Control (nonstress) Condition. The Ss were tested individually. At the beginning of the experiment, there was an 8-minute rest period during which S was required to complete the Zuckerman (1960) Affect Adjective Check List (AACL) by checking adjectives that described "how you feel at this moment." The AACL served as the measure of the phenomenological component of A-state.

Since Wechsler (1958) considered performance on Digits Backwards (DB) to be particularly sensitive to anxiety level, a special DB task was constructed for this study. After S completed the AACL, he was given DB from the WAIS (Wechsler, 1955, p. 41) in accordance with standard instructions to determine his limit, that is, the level at which he twice failed to repeat a series of a specified number of DB. He was then given two practice series, each one digit less than his limit, followed immediately by six more series of digits of the same length. The number of series correctly reproduced by S out of the six that were given to him was his Performance score.

After the Performance score was obtained, each S in the Failure Condition was told that, although he was not doing too badly, he could do better and that other Ss seemed to be doing better than he was. In the Control Condition, S was told that he was doing fine and that the next set of DBs were to be "just as easy as the ones he had just completed." Six more series of digits of the same length were then given to obtain a Test score for each S. Following the test series, S was required to fill out the AACL again, according to how he felt "while repeating the last few series of digits."

RESULTS

The effects of experimental conditions and level of A-trait on DS performance are presented in Figure 1. It may be noted that the

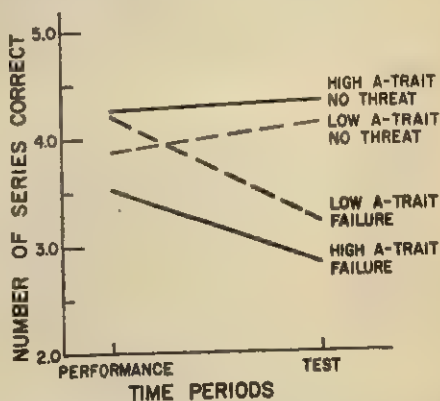


FIG. 1. The effects of failure and no-threat instructions and level of A-trait on digits backwards during the Performance and Test Periods ($N=18$ per group).

failure instructions produced decrements in DS performance. In an analysis of variance of these data, in which Conditions and A-trait level were the between-Ss variables and Time Periods the within-Ss variable, only the Conditions main effect ($F = 4.97$, $df = 1/68$, $p < .05$) and the Conditions \times Time Periods interaction effect ($F = 6.73$, $df = 2/68$, $p < .01$) were significant. The absence of any statistically significant effects involving the A-trait variable confirmed the impression gained from Figure 1 that DS Performance was unrelated to A-trait.

Next, the effects of experimental conditions and level of A-trait on intensity of A-state during the Test Period were evaluated. An examination of the distribution of A-state scores obtained during the Test Period indicated that the median AACL score was 12. The data for 14 Ss whose AACL scores fell at the median were eliminated; Ss with scores above and below the median were designated the high- and low-A-state groups, respectively. In order to provide an equal number of Ss in the high- and low-A-state groups, four Ss with AACL scores below the median were randomly eliminated.

As may be noted in Table 1, all of the high-A-trait Ss in the Failure group were classified as high A-state, and all but one of the low-A-trait Ss in the Control group were classified as low A-state. A statistical analysis of these data yielded a significant chi-square ($\chi^2 = 24.78$, $df = 3$, $p < .01$), indicating that the effect of experimental conditions on A-state was influenced by level of A-trait.

TABLE 1

FREQUENCY OF HIGH-A-TRAIT AND LOW-A-TRAIT
SUBJECTS IN THE FAILURE AND CONTROL
CONDITIONS WHO WERE CLASSIFIED AS
HIGH OR LOW IN A-STATE UNDER
STRESS

Conditions	Groups	N	Low A-state	High A-state
Control	Low A-trait	14	13	1
Control	High A-trait	13	6	7
Failure	Low A-trait	13	8	5
Failure	High A-trait	14	0	14
Totals		54	27	27

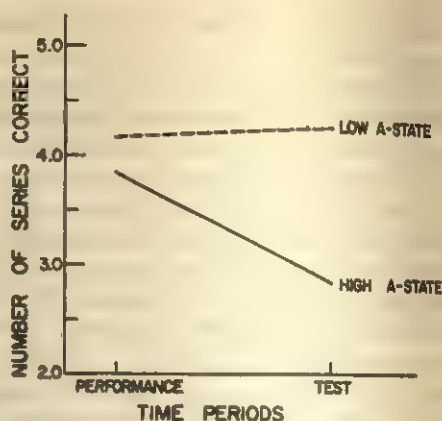


FIG. 2. The effects of low-A-state and high-A-state levels on Digits Backwards during the Performance and Test Periods ($N = 27$ per group).

The next step in the analysis of the data involved an evaluation of the effects of A-state on DS Performance. For this analysis, the data for the Failure and Control Conditions were combined. The mean DB scores obtained in the Performance and Test Periods by Ss in the high- and low-A-state groups are presented in Figure 2. These data were evaluated in an analysis of variance in which the most important finding was the significant A-state \times Time Periods interaction ($F = 6.52$, $df = 1/52$, $p < .01$). The main effects of A-state ($F = 5.50$, $df = 1/52$, $p < .05$) and Time Periods ($F = 4.90$, $df = 1/52$, $p < .05$) were also statistically significant. In order to evaluate further the effects of A-state on DB Performance, the data for each Time Period were separately analyzed by a Neuman-Keuls test (Winer, 1962), which indicated that the high- and low-A-state groups did not differ in the Performance Period ($F = .49$, $df = 1/52$) but that the DS Performance of the high-A-state group was significantly lower than that of the low-A-state group during the Test Period ($F = 11.05$, $df = 1/52$, $p < .01$).

DISCUSSION

In the present study, the failure to find a relationship between DS Performance and scores on the Taylor MA scale provides further evidence that decrements in DS are unrelated to measures of A-trait (Jackson & Bloomberg, 1958; Matarazzo, 1955; Walker

& Spence, 1964). The finding that Ss in the Failure Condition showed a significant decrement in DS performance was consistent with the results of previous investigations in which some form of stress was used to induce anxiety (Griffiths, 1958; Moldawsky & Moldawsky, 1952; Pyke & Agnew, 1963; Walker & Spence, 1964), and suggests that decrements in DS performance are associated with elevations in A-state.

Evaluating the effects of experimental conditions on level of A-state, it may be noted in Table 1 that high-A-trait Ss were particularly disposed to manifest high levels of A-state under conditions of failure-stress. It may also be noted that over half of the high-A-trait Ss responded with high levels of A-state in the Control Condition, even though they were told that they were doing well. In contrast, less than half of the low-A-trait Ss responded with high A-state in the Failure Condition and only one low-A-trait S responded with high A-state in the Control Condition. These results support the hypothesis that high-A-trait Ss tend to respond with high levels of A-state to situations that involve threats to self-esteem such as failure or threat of failure (Spielberger & Lushene, in press).

The finding of greatest theoretical importance in this study was the statistically significant decrement in DS performance for Ss who reported a high level of A-state. The fact that high-A-state Ss had lower DS scores than low-A-state Ss in the Test Period but not in the Performance Period is consistent with the interpretation that decrements in DS Performance are produced by elevations in A-state. This finding is also consistent with Trait-State Anxiety Theory (Spielberger & Lushene, in press), which posits that performance decrements on learning and memory tasks are produced by differential levels of A-state rather than differences in A-trait. According to this view, it would be expected that DS Performance will be related to A-trait only when differential levels of A-state are induced (by environmental circumstances or stressful experimental conditions) in Ss who differ in A-trait.

An alternative explanation for the decrements in DS performance that were observed in the present study is that Ss who did more poorly on the DS task experienced an increase in A-state in response to their deteriorating performance. This interpretation argues that the failure instructions in some way influenced performance without altering anxiety level, and that changes in A-state subsequently resulted from S's reaction to his poor performance. Although the experimental design in the present study does not permit a definitive choice between these alternative explanations of the findings, the first interpretation, which emphasizes the impact of anxiety level on performance rather than the reverse, would appear to be more consistent with the clinical literature.

REFERENCES

- CALVIN, A. D., KOONS, P. B., JR., BINGHAM, J. C., & FINK, H. H. A further investigation of the relationship between manifest anxiety and intelligence. *Journal of Consulting Psychology*, 1955, 19, 280-282.
- CATTELL, R. B., & SCHEIER, I. H. *The meaning and measurement of neuroticism and anxiety*. New York: Ronald Press, 1961.
- FREUD, S. *New introductory lectures in psychoanalysis*. New York: Norton, 1933.
- FREUD, S. *The problem of anxiety*. New York: Norton, 1936.
- GRIFFITHS, J. S. The effect of experimentally induced anxiety on certain subtests of the Wechsler-Bellevue. Unpublished doctoral dissertation, University of Kentucky, 1958.
- HODGES, W. F. Effects of ego threat and threat of pain on state anxiety. *Journal of Personality and Social Psychology*, 1968, 8, 364-372.
- JACKSON, D. N., & BLOOMBERG, R. Anxiety: Unitas or multiplex? *Journal of Consulting Psychology*, 1958, 22, 225-227.
- JURJEVICH, R. M. Interrelationships of anxiety indices of Wechsler intelligence scales and MMPI scales. *Journal of General Psychology*, 1963, 69, 135-142.
- LEWINSKI, R. J. The psychometric pattern: I. Anxiety neurosis. *Journal of Clinical Psychology*, 1945, 1, 214-221.
- MATARAZZO, R. G. The relationship of manifest anxiety to Wechsler-Bellevue subtest performance. *Journal of Consulting Psychology*, 1955, 19, 218.
- MERRILL, R. M., & HEATHERS, L. B. Deviations of Wechsler-Bellevue subtest scores from vocabulary level in university counseling-center clients. *Journal of Consulting Psychology*, 1952, 16, 469-472.
- MOLDAWSKY, S., & MOLDAWSKY, P. C. Digit span as an anxiety indicator. *Journal of Consulting Psychology*, 1952, 16, 115-118.

- PYKE, S., & AGNEW, N. McK. Digit span performance as a function of noxious stimulation. *Journal of Consulting Psychology*, 1963, 27, 281.
- RAPAPORT, D., GILL, M. M., & SCHAFER, R. *Diagnostic psychological testing*. Vol. 1. Chicago: Year Book Publishers, 1945. (Republished: rev. ed., New York: International Universities Press, 1968.)
- RAPAPORT, D., SCHAFER, R., & GILL, M. *Manual of diagnostic psychological testing. I: Diagnostic testing of intelligence and concept formation*. New York: Josiah Macy, Jr. Foundation, 1944.
- SPENCE, K. W. Anxiety (drive) level and performance in eyelid conditioning. *Psychological Bulletin*, 1964, 61, 129-139.
- SPIELBERGER, C. D. Theory and research on anxiety. In C. D. Spielberger (Ed.), *Anxiety and behavior*. New York: Academic Press, 1966.
- SPIELBERGER, C. D., & LUSHENE, R. E. Theory and measurement of anxiety states. In R. B. Cattell (Ed.), *Handbook of modern personality theory*. Chicago: Aldine, in press.
- TAYLOR, J. A. A personality scale of manifest anxiety. *Journal of Abnormal and Social Psychology*, 1953, 48, 285-290.
- WALKER, R. E., & SPENCE, J. T. Relationship between digit span and anxiety. *Journal of Consulting Psychology*, 1964, 28, 220-223.
- WARNER, S. J. The Wechsler-Bellevue psychometric pattern in anxiety neurosis. *Journal of Consulting Psychology*, 1950, 14, 297-304.
- WECHSLER, D. *Manual for the Wechsler Adult Intelligence Scale*. New York: Psychological Corporation, 1955.
- WECHSLER, D. *The measurement and appraisal of adult intelligence*. Baltimore: Williams & Wilkins, 1958.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.
- ZUCKERMAN, M. The development of an affect adjective check list for the measurement of anxiety. *Journal of Consulting Psychology*, 1960, 24, 457-462.

(Received July 22, 1968)

CLIENT-THERAPIST RELATIONSHIP FACTOR IN SYSTEMATIC DESENSITIZATION

JON E. KRAPFL

Drake University

AND

M. MIKE NAWAS¹

University of Missouri, Columbia

Forty snake-phobic female Ss were matched on degree of behavioral avoidance and subjective fear, and then randomly assigned to one of four groups: (a) a conventional desensitization group, using standard interpersonal procedures; (b) a semiautomated desensitization group in which treatment was delivered by tape recordings; (c) a pseudodesensitization control group; or (d) a no-treatment control group. The Ss in both desensitization groups showed significantly greater improvement than did Ss in either of the two control groups. No differences were found between conventional and semiautomated procedures. It was concluded that the relationship is neither a primary nor critical factor in desensitization.

In a review of the psychotherapy of phobias, Andrews (1966) concludes that the central factor in the treatment of phobias is a therapist who establishes himself in a direct nurturant role and uses this relationship as leverage to encourage the patient to confront fear-arousing situations. Andrews maintains that the behavior therapists' informal and casual discussions of their interchanges with their patients reveal a rather characteristic type of social relationship which plays an important role in the modification of behavior.

Meyer and Gelder (1963) express a point of view similar to that of Andrews. They, too, emphasize the importance of the relationship between the patient and the behavior therapist, and even suggest that the relationship is the single factor which seems most relevant to successful outcome of behavior therapy.

The patient-therapist relationship and its effects on the outcome of behavior therapy is an important empirical issue which deserves direct and critical exploration. Beginnings have been made by Lang, Lazovik, and Reynolds (1965), Melamed and Lang (1967), and Paul (1965). Altogether, these studies not only disallow an unequivocal answer to the problem under consideration, but they impart further good reasons for conducting experiments expressly designed to shed light on the significance of the relationship in behavior therapy. In the study by Paul, for

example, while the attention-placebo therapy group did not change as much as the desensitization group it did show greater improvement than a no-treatment control group. Melamed and Lang (1967), comparing the performance of a therapist with that of an apparatus programmed for the reduction of fear, found the automated procedure about as effective as the live therapist. In their research, however, the human *E* was actively involved in some critical aspects of the automated program; he conducted the first training sessions in hierarchy building and the initial relaxation instruction sessions.

The present study was specifically designed to assess the relationship factor, using controls employed in previous investigations as well as controls for several factors which have either been neglected or inadequately met in a number of previous experimental studies. More specifically, all instructions to all Ss were delivered by tape recordings; Ss in the various groups were matched on their degree of avoidance behavior to the phobic stimulus (live snake) in an actual behavioral situation as well as on their subjective ratings of fear. Finally, in order to make possible more rigorous comparisons between groups, two modifications in the standard desensitization procedure were introduced. First, no detailed history was taken from each *S*, and second, a standard rather than an individualized hierarchy was used for all Ss. These modifications—history taking and type of hierarchy—have been shown to have no un-

¹ Requests for reprints should be sent to M. Mike Nawas, Department of Psychology, University of Missouri, Columbia, Missouri 65201.

toward effects on the efficacy of desensitization (Emery & Krumboltz, 1967).

METHOD

Measures

Measures taken in this experiment were those used in the initial screening and selection of Ss, and those employed in testing the effects of experimental procedures.

A Preliminary Screening Questionnaire (PSQ) was administered to all undergraduate students taking the introductory psychology course at the University of Missouri at Columbia. These students are required to participate in one or more studies, of their own choosing, to satisfy course requirements. The PSQ consisted of two main parts, one pertaining to fears of animals and the other to the individual's ability to closely follow imaginary scenes suggested by another person.

The S was asked to imagine himself in a room 15 feet in length at the end of which was a caged animal, and to indicate if he could (a) approach the animal, (b) touch it with a gloved hand, or (c) touch it with a bare hand. The S was asked to indicate whether he could perform the above steps for each of six animals: rat, nonpoisonous snake, spider, mouse, cockroach, and beetle. Because more Ss indicated that they were afraid of snakes than of any of the other six animals, the fear of nonpoisonous snakes was selected for study. The Ss who indicated that they would be able to touch the snake with a gloved or bare hand were not considered sufficiently fearful of snakes and were, therefore, excluded from participation in the experiment.

The PSQ also contained the "imagery test." This was adapted from experimental work in hypnosis. The S was asked to imagine in detail the last meal he ate. Questions were then asked concerning (a) the clarity of the image, (b) whether or not it was in color, and (c) whether or not S was able to project himself into the scene which involved the meal in question. Those who did not answer all three questions in the affirmative were excluded from the study.

The Behavioral Avoidance Test (BAT), similar to that used in previous desensitization studies with animal phobic Ss (Lang, Lazovik, & Reynolds, 1965), was used to measure the intensity of the individual's avoidance response to the feared object. The test was conducted in a room 6 × 18 feet containing only a table next to the wall opposite the entrance. A white cage (24 × 18 × 12 inches), which opened from the top, was placed on the table, clearly exposing the 36-inch nonpoisonous Speckled-back King Snake through a plate-glass front. The test itself consisted of a check list of 20 separate items corresponding to 20 discreet approach steps toward the snake, ranked in order of increasing aversiveness. The steps ranged from (1) standing outside the door of the testing room, to (20) picking up the snake and bringing it close to the chest.

Each S was escorted individually into an anteroom which opened into the testing room, was seated, and asked to listen to recorded instructions prepared in advance by E. The recording informed S of the snake in the next room and he was told that he was to attempt to perform, in order, each step in the avoidance check list. The S was alone in the room which housed the snake but was observed through a one-way mirror in an adjoining room. The Ss who were able to reach Step 14 on the check list (placing their hand into the cage) were excused from the study.

In addition to S's behavioral response, a measure of subjective fear—Fear Thermometer (FT)—was also obtained. The FT has been rather extensively employed in previous studies, for example, Lazovik and Lang (1960), with inconclusive results. The measure consists of a 10-point scale ranging from "calm" to "terrified" on which S is asked to indicate the intensity of his fear at the point beyond which he could proceed no further in the BAT.

The Fear Survey Schedule II (FSS-II), developed by Geer (1965), was used for having Ss rate the degree of their fear for each of 50 objects or situations. The total score was taken as S's measure of generalized fear. A separate analysis was performed on Item 38 of the FSS-II, the snake item, to determine whether experimental procedures specifically affected S's response to that item.

Subjects

The Ss in the study were 40 female undergraduates, selected from the initial pool of approximately 1,200 students who were administered the PSQ described earlier. Of these, 308 Ss indicated that they were fearful of snakes. Eleven males and 15 females were rejected for failure to pass the imagery test. An additional 84 females and 19 males were rejected for failing to display sufficient avoidance behavior in the presence of the snake. Finally, one male and five females were excused from participation in the experiment because they had had or were then engaged in psychiatric treatment. Due to the relatively small number of males remaining eligible it was decided that only female Ss would be used for the study. Each of the 130 female Ss who remained eligible was assigned to 1 of 13 groups. This study is concerned with data relevant to four of these groups.

Procedure

The 40 Ss were matched on the basis of their BAT and FT scores. Assignment of the matched Ss to one of the groups was then done on a random basis. Identification of each of the four groups follows:

Live desensitization (LD). Each S in this group received one 30-minute taped training session in relaxation followed by five 30-minute sessions of systematic desensitization. The desensitization sessions were conducted by E using standard interpersonal procedures as described by Wolpe (1958).

Taped desensitization (TD). Each *S* in this group received one taped 30-minute training session in relaxation followed by five 30-minute taped sessions of systematic desensitization.

Pseudodesensitization (PD). The *Ss* in the *PD* group received one taped 30-minute training session in relaxation followed by five 30-minute sessions of pseudodesensitization. The imaginal stimuli used for this group were pleasant and unrelated to the phobic stimulus. The purpose of this group was to control for such factors as expectancy of improvement, suggestion, and attention.

No treatment (NT). The *Ss* in this group received no treatment and were given only the pretreatment, posttreatment, and follow-up test batteries.

The apparatus used during treatment sessions consisted of a tape recorder, two sets of headphones, and a black vinyl reclining chair. For all *Ss* the *E* was present in the room. For *Ss* in the *LD* condition the procedure was to escort *S* into the room, seat her facing *E*, and engage in conversation. For *Ss* in the *TD* condition *E* was seated out of sight of *S* and did not communicate with her during the entire session. Desensitization was administered by means of tape recordings and was received by *S* through a set of headphones. The headphones were set up to allow *E* to monitor *S's* responses as both *E* and *S* listened to the recordings.

Treatment sessions were conducted in a different room from the one in which testing was held. Each *S* who received relaxation training was seated in the lounge chair and told that everything would be explained to her by means of a tape recorder. The taped relaxation instructions were similar to those used by Paul (1965) and others. Desensitization sessions began with a 3-minute relaxation induction sequence followed by the desensitization itself. Each item in the hierarchy of 20 steps (of the BAT) was presented until *S* was able to entertain an image of the item for 15 seconds on two consecutive occasions without signaling anxiety.

The posttreatment battery was administered to each *S* immediately upon her completion of the fifth treatment session. At that time measures of *S's* responses were again taken on the BAT, the FT, and the FSS-II, in that order. In the follow-up session, 6 weeks after the completion of the final treatment session, the BAT, FT, and FSS-II were administered again.

RESULTS

The results are reported separately for each measure used. Separate analyses were performed on the data of the BAT, the FT, the FSS-II, and the FSS-II snake item. Pretest data were analyzed on the basis of raw scores, whereas the posttest and follow-up data were analyzed in terms of change scores between pretest and posttest, then between pretest and follow-up tests, respectively.

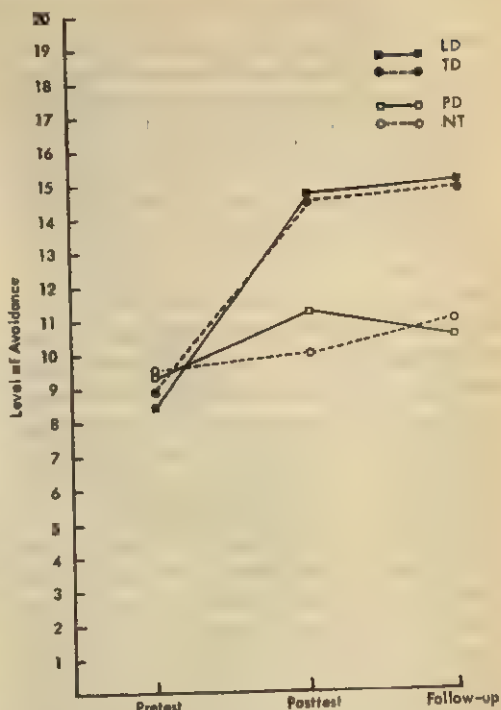


FIG. 1. Mean pretest, posttest, and follow-up test scores on the BAT.

The *Ss* were matched on their pretest BAT scores, so there were, of course, no significant differences between the groups and the pretest analysis ($F < 1$). The analysis of variance on posttest change scores, however, did reveal significant differences ($F = 7.30$, $p < .001$). The overall strength of the statistical association (ω^2) was estimated at .35, indicating that the treatment procedures accounted for a sizable portion of the systematic variation. Individual comparisons of posttest BAT scores reveal that both the *LD* and the *TD* groups showed significantly more

TABLE 1
TUKEY'S HSD PRETEST-TO-POSTTEST COMPARISONS
ON THE BEHAVIORAL AVOIDANCE TEST

Group	NT	PD	TD	LD
NT	—	1.5	5.1**	5.3**
PD	—	—	3.6*	3.8*
TD	—	—	—	.2

* $p < .005$.

** $p < .001$.

TABLE 2

TUKEY'S HSD PRETEST-TO-FOLLOW-UP COMPARISONS
ON THE BEHAVIORAL AVOIDANCE TEST

Group	NT	PD	TD	LD
NT	—	—	4.7*	4.3*
PD	—	—	5.0*	4.6*
TD	—	—	—	.4

* $p < .001$.

approach behavior in the BAT posttest session than did the NT or the PD controls. An analysis of variance on the follow-up scores indicated that the group differences were maintained over the 6-week follow-up period ($F = 7.19$, $p < .001$). The ω^2 was .36, virtually the same as in the posttest analysis.

Individual comparisons of group means show that the differences found on the posttest analysis were essentially maintained. Tukey's HSD comparisons (Winer, 1962) of pretest to posttest and follow-up BAT scores are shown in Tables 1 and 2, respectively. The degree of avoidance of the snake in the pretest, posttest, and follow-up sessions is depicted in Figure 1.

No significant differences between the groups existed on the FSS-II snake item prior to treatment ($F < 1.33$). However, there were differences in the amount of snake fear reduction reported by the groups following

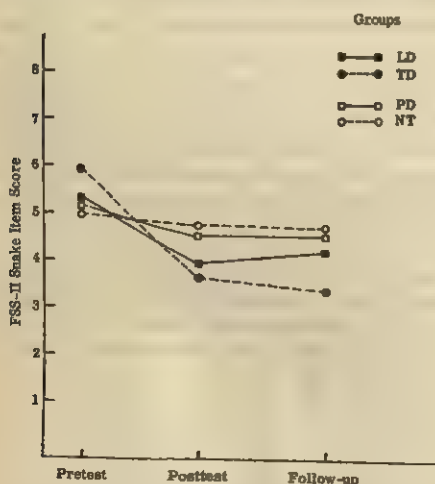


FIG. 2. Mean pretest, posttest, and follow-up test scores on the FSS-II snake item.

TABLE 3

TUKEY'S HSD PRETEST-TO-POSTTEST COMPARISONS
ON THE FSS-II SNAKE ITEM

Group	NT	PD	LD	TD
NT	—	.3	1.8*	2.5***
PD	—	—	1.5*	2.2**
LD	—	—	—	.7

* $p < .05$.
 ** $p < .005$.
 *** $p < .001$.

treatment ($F = 4.61$, $p < .01$). The strength of statistical association (ω^2) here was .23. Both the LD and TD groups reported significantly less fear of snakes than did either the NT control or the PD group. Again, the NT and PD controls were not significantly different from one another on this measure.

Essentially the same results were found on the follow-up analysis ($F = 4.23$, $p < .01$). The strength of statistical association (ω^2) was again .23. Tukey's HSD comparisons of pretest to posttest and follow-up FSS-II snake-item scores are shown in Tables 3 and 4, respectively. Figure 2 illustrates the means of self-rated fear of snakes for the various groups.

Comparison of the groups on the FT as well as on the overall FSS-II revealed no significant difference before treatment, from pretreatment to posttreatment, or from pretreatment to follow-up.

DISCUSSION

Concerning the fundamental question of this study, that is, the possible effect of the relationship factor on outcome, virtually no differences between the LD and TD groups were found on any of the measures used. Only

TABLE 4

TUKEY'S HSD PRETEST-TO-FOLLOW-UP COMPARISONS
ON THE FSS-II SNAKE ITEM

Group	PD	NT	LD	TD
PD	—	.1	1.4*	2.3**
NT	—	—	1.3*	2.2**
LD	—	—	—	.9

* $p < .05$.
 ** $p < .001$.

.2 points separated these groups on their pretest to posttest change scores on the BAT. On the follow-up the difference was .4 points. Also, minute differences between these two groups were found in the other measures used in the study. These results support the work of Melamed and Lang (1967), in which they found no differences between automated desensitization and desensitization conducted in the standard fashion by a therapist.

It might be argued that, even though the entire treatment for the TD group was conducted by means of tape recordings and Ss did not interact with *E* during the treatment procedures, some sort of relationship was established because *E* was present in the room. While it is true that the procedure was not fully automated, these criticisms seem to be begging the question somewhat. No attempt to establish a relationship between *E* and *S* in the TD group was made and, indeed, the impersonal nature of their interaction was more conducive to distancing than to relating. There was virtually no communication and *E* deliberately tried to avoid interpersonal involvement with Ss who received the semi-automated treatment.

The proposition put forth by Andrews (1966) and others that the therapist must accept the patient, establish a nurturant role, and then use the relationship as leverage to make the patient confront his fears does not seem to be supported by the results of this study.

This is not to say, of course, that the therapeutic relationships have no effect. It has been demonstrated that clients can associate the therapist with the discriminative properties of the reinforcing agent; and the therapist can use these qualities to bring about changes in the client's behavior. The point, however, is that the relationship factor, as demonstrated in this experiment—which is merely an analogue to the clinical helping relationship—is not an integral and necessary part of the desensitization process, and successes with the technique cannot be attributed solely or even primarily to this factor.

Two side observations are worth noting. First, to the extent that the control of modeling, suggestion, attention, and *E* bias through the use of taped instructions was attained in this study, the results strengthen still further the credibility of previous outcome studies in desensitization. Second, despite the modifications of the systematic desensitization procedure introduced in this study, the treatment appears to have been at least as effective as that obtained in previous experimental studies. It may be speculated that the recording of thorough case histories, the construction of separate hierarchies, and conducting more than five sessions of desensitization are often unnecessary, at least insofar as the treatment of well-defined fears is concerned.

REFERENCES

- ANDREWS, J. D. Psychotherapy of phobias. *Psychological Bulletin*, 1966, 66, 455-480.
- EMERY, J. R., & KRUMBOLTZ, J. D. Standard versus individualized hierarchies in desensitization to reduce test anxiety. *Journal of Counseling Psychology*, 1967, 14, 204-209.
- GEER, J. The development of a scale to measure fear. *Behavior Research and Therapy*, 1965, 3, 45-53.
- LANG, P. J., LAZOVIC, A. D., & REYNOLDS, D. J. Desensitization, suggestibility, and pseudotherapy. *Journal of Abnormal and Social Psychology*, 1965, 70, 395-402.
- LAZOVIC, A. D., & LANG, P. J. A laboratory demonstration of systematic desensitization psychotherapy. *Journal of Psychological Studies*, 1960, 11, 238-247.
- MELAMED, B., & LANG, P. J. Study of the automated desensitization of fear. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, May 1967.
- MEYER, V., & GELDER, M. G. Behavior therapy and phobic disorders. *British Journal of Psychiatry*, 1963, 109, 19-28.
- PAUL, G. L. *Insight versus desensitization in psychotherapy: An experiment in anxiety reduction*. Stanford: Stanford University Press, 1965.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.
- WOLPE, J. *Psychotherapy by reciprocal inhibition*. Stanford: Stanford University Press, 1958.

(Received July 26, 1968)

EFFECTS OF VICARIOUS THERAPY PRETRAINING AND ALTERNATE SESSIONS ON OUTCOME IN GROUP PSYCHOTHERAPY WITH OUTPATIENTS¹

CHARLES B. TRUAX²
University of Florida

AND

DONALD G. WARGO
*Arkansas Rehabilitation Research and Training
Center, University of Arkansas*

The present study investigated the effects of therapist-absent alternate sessions and a vicarious therapy pretraining experience on group psychotherapy outcome with outpatients. It also related patient self-exploration to outcome. The alternate sessions regimen and high levels of self-exploration both led to slightly better outcome. Vicarious therapy pretraining was found to be highly facilitative, having its greatest effect on neurotic symptomatology.

The use of alternate sessions in group psychotherapy has led to some controversy regarding its efficacy. By alternate sessions is meant the use of group therapy sessions in which the therapist is absent, alternating with sessions in which both patients and therapist are present. Their use has been recommended by Wolf (1961), Wolf and Schwartz (1962), and Truax (1962a, 1962b). On the other hand, Slavson (1963) and Ginott³ have raised objections to the use of alternate sessions. It is quite apparent that, if helpful, alternate sessions would constitute an economically desirable regimen. Previous research by Truax and Carkhuff (1965) on a sample of hospitalized patients indicated that the use of alternate sessions was not only nonbeneficial, but led to significantly poorer outcome than regular group psychotherapy, supporting the position taken by Slavson and Ginott. However, since outpatients as a group are in closer contact with reality and more adequate in interpersonal relating, it seems plausible that alternate sessions might prove beneficial with neurotic outpatients even though they might be harmful with hospitalized patients.

Vicarious therapy pretraining (VTP) has been proposed by Truax (1962a, 1962b) to provide standard cognitive and experiential structuring of "how to be a good patient" as

a means of quickly engaging the patient in the process of group therapy. VTP simply involves presenting to prospective patients a 30-minute tape recording of excerpts of "good" patient in-therapy behavior. It provides a vicarious experience of how clients often explore their problems and feelings, and how they prove helpful to one another prior to their introduction to group therapy. One study (Truax & Carkhuff, 1965) has indicated modest therapeutic benefit of VTP.

Prior research in individual and group psychotherapy has also suggested that successful patients engage in a greater degree of self-exploration than less successful patients in a wide variety of patient populations (Blau, 1953; Braaten, 1958; Peres, 1947; Seeman, 1949; Tomlinson & Hart, 1962; Truax, 1961, 1962a, 1962b; Truax & Carkhuff, 1963; Wagstaff, Rice, & Butler, 1960). The present study attempted, also, to study the relationship between patient in-therapy self-exploration as it is related to outcome.

The present study, involving a heterogeneous sample of both therapists and patients, was aimed at evaluating the following hypotheses: (a) patients receiving alternate sessions in addition to their regular sessions will show greater evidence of constructive personality change than patients receiving regular sessions only, (b) patients receiving VTP prior to group therapy will show greater evidence of constructive personality change than patients not receiving VTP, and (c) patients in groups that engage in relatively high levels of self-exploration will show greater evidence of constructive personality change than pa-

¹ This research was supported by Rehabilitation Services Administration Grant RD-906 and Social and Rehabilitation Services Grant RT-13.

² Requests for reprints should be sent to Charles B. Truax, College of Education, University of Florida, Gainesville, Florida 32601.

³ H. G. Ginott, personal communication, March, 1964.

tients in groups that engage in relatively low levels of self-exploration.

METHOD

A total of eight psychotherapy groups, with an initial 10 patients in each group ($N=80$), were seen by eight different therapists⁴ for time-limited group psychotherapy consisting of 24 sessions on a twice weekly basis over a time span of approximately 3 months. The 80 outpatients were assigned to groups from the available populations of the following co-operating institutions: Department of Psychiatry, University of Wisconsin; Counseling Center, University of Massachusetts; Veterans Administration Clinic, Covington, Kentucky; and University of Kentucky.

All group therapy sessions were completely tape recorded to allow for analysis of levels of patient self-exploration occurring in each group. Patients were given a battery of psychological tests pretherapy and posttherapy which served as the basic measures of outcome.

The eight groups formed a 2×2 factorial design, with half of the groups receiving VTP and the other half non-VTP, and half of the groups receiving alternate sessions in addition to regular sessions with the other half holding regular sessions only. For those patients receiving VTP, the 30-minute VTP tape recording was presented to the group prior to the first session. The alternate sessions for those groups holding alternate sessions began after the eighth, ninth, or tenth sessions, and continued alternating with the regular sessions until the 24 regular sessions were completed.

Patient Population

The patient population as a whole was a mildly disturbed, neurotic outpatient group. The patients ranged in age from 18 to 58, with a distribution skewed to the right. Six patients, based on their pretesting, might be labeled schizophrenics.

Therapists

The therapists were assigned to the eight groups within the samples from the participating institutions. In general, they were highly experienced in individual and group psychotherapy. They were quite heterogeneous in orientation, including two therapists who could be described as client-centered, four therapists who could be described as psychoanalytically oriented, and two who would describe themselves as eclectic in orientation.⁵ Of the thera-

pists and co-therapists, four were psychiatrists and seven were psychologists.

The therapists were, in general, positively disposed toward the study. They were optimistic about the usefulness of VTP, and they were interested in the effects of alternate sessions. The latter constituted a minor administrative annoyance, but there was no evidence that therapists felt threatened by their groups' meeting without them.

Measurement of Patient Self-Exploration

The Depth of Intrapersonal (Self) Exploration Scale (Truax, 1962c) was used as the basic measure of patient self-exploration. From the set of tape recordings of regular sessions of the eight groups, one 3-minute sample was selected from the middle third of every other recorded therapy session (89 samples). These samples were obtained by running the tape recorder at rewind to the middle section of the tape and stopping. The 3-minute samples were then taken beginning with the first new utterance by a patient or therapist and continuing for 3 minutes, or until the conclusion of a statement to preserve intelligibility. These samples were rerecorded individually onto small single spools which were then randomly assigned code numbers and presented to the raters in boxes of 12 samples.

Four undergraduate college students who were naive with respect to psychotherapy theory and practice were trained in the use of the rating scale. A criterion of a correlation of .50 for interjudge and rate-rater reliability on training samples was used to screen the raters. The raters rated the coded samples in different orders of blocks of 12 samples. The mean ratings from the four judges per sample were used in the analyses.

Measurement of Patient Personality Change

Although 80 patients were given pretesting, a number of patients dropped out of group therapy before the tenth session, some refused to participate in testing posttherapy, and for a few the test administrations were judged invalid and discarded without examination of the results because of patient misunderstanding or inability to perform on the measure. Thus, 62 patients were available for data analysis on some measures, while as few as 51 patients were available for analysis on the measures of Q sorts for self-concepts.

The test battery in the present study administered pretherapy and posttherapy included the MMPI, with the 10 clinical scales, the Sum of Clinical scales, the Barron Ego Strength scale, the Truax Constructive Personality Change Index, the Welsh Internalization of Anxiety Ratio, Edwards Social Desirability, Sum of Validity Deviations, Welsh's Anxiety Index; the Q sort for self- and ideal concept (Butler & Haigh, 1954) with the five specific measures of self- and ideal concepts; and the Finney Palo Alto Group Therapy scale (Finney, 1954).

The 80-item Q -sort deck was used for only two sorts (self and ideal) both pretherapy and post-

⁴ Several groups involved co-therapists in addition to the therapist responsible for the group.

⁵ The present research would not have been possible without the dedicated contribution of the therapists who gave so freely of their time, energy, and talent. Appreciation is very gratefully extended to the following therapists and co-therapists: Arnold Marx, L. Stein, Donald Kiesler, Donald Price, Robert R. Carkhuff, Jerome Sczymanski, Walter A. Dickenson, Joseph Willett, Dean Allen, and Joe Havens.

therapy. The five measures obtained were (a) change in the correlation between self and expert *Q* sorts from pretherapy to posttherapy, (b) change in the correlation between self and ideal *Q* sort, (c) change in the correlation between ideal and expert *Q* sort, (d) change in the self *Q* sort adjustment scores, and (e) change in ideal adjustment scores. The "expert" *Q* sort is a *Q* sort representing ideal adjustment made by a panel of 14 experts (Lewis, 1959). The "adjustment" *Q* sorts, both self and ideal, refer to a method of scoring the *Q* sort on the basis of an imposed true-false type of dichotomy.

The Palo Alto Group Therapy scale consists of a number of true-false items referring to social interaction behavior engaged in by the patient during group psychotherapy itself. It is filled out by the therapist. In the Forsythe and Fairweather (1961) study it was the only measure in a large battery that was predictive of follow-up adjustment post-therapy.

For the most part the outcome battery consisted of self-report measures, however. Ideally, behavioral indexes of change might have been included, such as ratings by *Os*, changes in earnings, grades, etc.; but the heterogeneity of the patient population and geographic dispersion of the groups made this infeasible.

RESULTS

Change scores for each patient on each of the 23 measures of outcome were obtained by subtracting the score earned pretherapy from the score earned posttherapy. To guard against possible differences between groups in initial status, analysis of covariance was used to partial out the pretherapy level in its effect on measured change.

The basic design of the present study might be described as a 2×2 factorial design with the additional factor of high versus low self-exploration within each cell. Perfect design balance was disturbed, however, by missing data due to invalid testing, refusal to take tests, and premature terminations. As such, cell frequencies were unequal requiring that the unweighted means method for computing the analyses of covariance be used. In addition, since cell frequencies become more disproportionate with finer subdivisions, and since presumably the greater the disproportionality the more questionable the validity of the analysis, the major variables were analyzed separately by one-way analyses of covariance. None of the hypotheses in the study are concerned with interactions of the vari-

ables, thus lack of information regarding these is not a great loss. To test the main effects in each analysis, group or therapist variability within cells was used as the estimate of error. This constitutes a relatively conservative error estimate. Differences in outcome associated with different therapists and therapy groups are assessed by testing between group variability against *S* variability.

Effects of Vicarious Therapy Pretraining

The mean values of outcome measures for patients receiving VTP versus those not receiving VTP (NVTP) is shown in Table 1. There is a striking tendency for VTP to show therapeutic advantage over NVTP. On the 23 measures of outcome, patients receiving VTP showed greater improvement on 21 and less improvement on only two ($\chi^2 = 15.696$, $p < .001$). Also presented in Table 1 are the results of the analyses of covariance for the effects of VTP on patient outcome. Patients receiving VTP showed significantly greater improvement than patients not receiving VTP on the following MMPI Scales: *D*, *Hy*, *Pt*, and the Sum of Clinical Deviations. Thus, the data tend to provide relatively strong support for the initial hypothesis. It appears that the greatest influence of VTP was on neurotic symptomatology.

Effects of Alternate Sessions

Table 2 presents the mean values on the 23 outcome measures for patients receiving alternate sessions in addition to regular sessions versus patients receiving only regular sessions. Patients receiving alternate sessions showed greater improvement on 17 measures and less improvement on six ($\chi^2 = 2.630$, *ns*); whereas, patients receiving only regular sessions showed greater improvement on only six measures and less improvement on 17.

The results of the analyses of covariance on the 23 measures also appear in Table 2. There it can be seen that the addition of alternate sessions to regular group therapy results in significantly greater outcome on the following MMPI measures: *Hy*, Sum of Clinical Deviations, and the Constructive Personality Change Index.

Thus, the data suggest moderately positive consequences by the addition of alternate sessions to regular outpatient group meetings.

Patient Degree of Self-Exploration

The level of self-exploration was calculated for each group, and then the groups were subdivided into those showing highest versus those showing lowest levels of self-exploration. Ratings were made with raters using the Scale of Depth of Self-Exploration (Truax, 1962c).

The mean outcome scores on the 23 measures for patients in groups that engaged in high levels of self-exploration and patients in groups engaging in low levels of self-exploration are presented in Table 3. There it can be seen that patients in groups engaging in high levels of self-exploration showed

greater improvement on 15 measures and less improvement on eight ($\chi^2 = 2.130$, *ns*); whereas, for patients in groups engaging in low levels of self-exploration, greater improvement occurred on only eight measures and less improvement on 15.

The results of the analyses of covariance for high and low levels of self-exploration are also presented in Table 3. Differences favoring better outcome with high rather than low levels of self-exploration occurred on the following measures: *Q* sort for self-adjustment, *Q* sort congruence between self- and ideal concepts, and the Constructive Personality Change Index from the MMPI. These findings give modest support to the hypothesis suggesting greater improvement for high self-exploration groups.

TABLE 1
MEAN VALUES AND ANALYSIS OF COVARIANCE *F* RATIOS: EFFECTS OF VICARIOUS
THERAPY PRETRAINING ON OUTCOME

Measures of outcome	Mean values		Analysis of covariance <i>F</i> ratios	
	VTP	NVTP	VTP-NVTP	Therapists/Cells
<i>Q</i> sort				
Self, adjustment	9.950	4.065		32.745***
Self-expert <i>r</i>	.335	.138	1.187	34.256***
Ideal, adjustment	.952	-.871		53.533***
Self-ideal <i>r</i>	.354	.177		27.884***
Ideal-expert <i>r</i>	.001	.003		45.896***
Alto Group Therapy scale	.490	.179	4.019	1.475
MMPI				
Sum of Clinical Deviations	23.792	-.750	8.034*	
Constructive Personality Change Index	14.917	4.125		
<i>Hs</i>	1.125	-1.188	2.033	
<i>D</i>	4.000	-.125	5.758*	
<i>Hy</i>	3.375	-1.188	19.631**	
<i>Pd</i>	2.042	-.906	2.958	
<i>Mf</i>	.792	1.406		1.227
<i>Pa</i>	2.792	.469		
<i>Pt</i>	4.875	.313	5.554*	
<i>Sc</i>	6.083	1.250	1.779	
<i>Mc</i>	1.583	.781		4.191***
<i>Si</i>	5.500	2.531		
Barron's Ego Strength	3.667	1.500		
Edwards' Social Desirability	3.750	2.500		1.124
Sum of Validity Scales' Deviations	3.875	-.938	2.010	1.072
Welsh's Anxiety Index	7.125	1.906	2.944	
Welsh's Internalization Ratio	.025	.016		1.394

* $p < .05$, one-tailed test.
 ** $p < .05$, two-tailed test.
 *** $p < .01$, two-tailed test.

DISCUSSION

In general, the results of the present study lend support to the hypotheses stated at the outset. As such, the study was successful in serving to establish the effectiveness of the variables investigated in producing constructive personality and behavioral change through the treatment known as group psychotherapy.

The effect of inserting alternate sessions with the therapist absent between regular group therapy sessions with the therapist present would appear to be slightly facilitative from the present results. They generally support the recommendations of Wolf (1961) and Wolf and Schwartz (1962), and refute the suggestions of Slavson (1963) and Ginott (see Footnote 3) who have objected to the use

of alternate sessions. This finding is of particular importance as the shortage of trained therapists increases. It is an economical procedure since it extends the effectiveness of scarce therapists. It should be cautioned, however, that alternate sessions may not be effective with other types of patient populations. Indeed, it has been found by the authors (Truax & Wargo, 1968a, 1968b) that the use of alternate sessions with hospitalized mental patient groups and with juvenile delinquent groups led to considerably less improvement than did the use of regular sessions only. It may be that patients must be more socially responsible in order to benefit from alternate sessions; indeed, in order to be not harmed or retarded by alternate sessions.

TABLE 2
MEAN VALUES AND ANALYSIS OF COVARIANCE *F* RATIOS: EFFECTS OF
ALTERNATE SESSIONS ON OUTCOME

Measures of outcome	Mean values		Analysis of covariance <i>F</i> ratios	
	Alternate	Regular	Alternate regular	Therapists/ Cells
<i>Q</i> sort	8.000	4.048		32.745***
Self, adjustment	.260	.151		34.256***
Self-expert <i>r</i>	-.226	0.000		37.848***
Ideal, adjustment	.324	.125		27.884***
Self-ideal <i>r</i>	.025	-.032		40.761***
Ideal-expert <i>r</i>	.254	.421		1.475
Palo Alto Group Therapy Scale				
MMPI				
Sum of Clinical Deviations	15.290	2.920	11.517**	
Constructive Personality Change Index	12.419	4.200	21.304**	
<i>Hs</i>	1.032	-1.720	7.049*	
<i>D</i>	1.903	1.320		
<i>Hy</i>	2.194	-1.000	23.572**	
<i>Pd</i>	.387	.320		1.227
<i>Mf</i>	.452	2.000		
<i>Pa</i>	1.645	1.240	1.244	
<i>Pt</i>	3.290	1.000	2.923	
<i>Sc</i>	4.290	2.120	2.994	
<i>Ma</i>	.323	2.120		4.191***
<i>Si</i>	4.484	2.960		
Barron's Ego Strength	3.323	1.320	4.210	1.124
Edwards' Social Desirability	2.903	3.200		
Sum of Validity Scales' Deviations	-.161	2.720	1.712	
Welsh's Anxiety Index	3.677	4.720		1.394
Welsh's Internalization Ratio	.043	-.008		

* $p < .05$, one-tailed test.

** $p < .05$, two-tailed test.

*** $p < .01$, two-tailed test.

The effects of vicarious therapy pretraining were quite positive. This supports the hypothesis based on earlier suggestions by Truax (1962a, 1962b) that this form of systematic preparation for psychotherapy might well be beneficial to outcome. Again, however, this finding should not be generalized to other patient populations since this same hypothesis is clearly not supported by the findings in the previously mentioned study of institutionalized juvenile delinquents. It does appear, however, that the structuring of patient roles through the use of tape-recorded modeling is facilitative with psychoneurotic outpatients. This procedure is of therapeutic value when incorporated into the treatment of neurotic outpatient groups.

As for the relationship between level of self-exploration and psychotherapeutic outcome, the results of the present study provide modest support for the hypothesis that high levels of self-exploration will be related to high degrees of outcome. It is of particular interest to find that it holds for group therapy, since level of self-exploration in group therapy is determined by the overall level of the group rather than by the levels of self-exploration demonstrated by individual members of the group. It is most likely that there is considerable within-group heterogeneity of self-exploration in most groups. As such, measurement of the variable is gross. At the same time, it is likely that some benefit accrues to the patient who sits quietly but

TABLE 3
MEAN VALUES AND ANALYSIS OF COVARIANCE *F* RATIOS: EFFECTS OF DEPTH OF
INTRAPERSONAL EXPLORATION ON OUTCOME

Measures of outcome	Mean values		Analysis of covariance <i>F</i> ratios	
	High DX	Low DX	High-Low DX	Therapists/ Cells
<i>Q</i> sort				
Self, adjustment	9.233	2.286	4.553*	19.506***
Self-expert <i>r</i>	.293	.104	3.510	21.961***
Ideal, adjustment	-.290	.095		38.288***
Self-ideal <i>r</i>	.330	.116	6.336*	16.556***
Ideal-expert <i>r</i>	-.024	.041	2.255	30.675***
Palo Alto Group Therapy Scale	.283	.386	1.485	3.233***
MMPI				
Sum of Clinical Deviations	12.533	6.577		1.207
Constructive Personality Change Index	15.833	.577	4.950*	
<i>Hs</i>	-.267	-.115		1.188
<i>D</i>	1.967	1.269		1.432
<i>Hy</i>	1.133	.346		
<i>Pd</i>	.233	.500		
<i>Pd</i>	.633	1.731	3.377	
<i>Mf</i>	2.067	.769		
<i>Pa</i>	3.333	1.038		1.017
<i>Pt</i>	4.367	2.115		1.049
<i>Sc</i>	.167	2.231	1.677	2.577**
<i>Ma</i>	5.833	1.462	3.570	
<i>Si</i>	3.600	1.077	3.113	
Barron's Ego Strength	4.533	1.308	2.729	
Edwards' Social Desirability	.833	1.462		1.321
Sum of Validity Scales' Deviations	5.500	2.577		
Welsh's Anxiety Index	.050	-.015	2.451	
Welsh's Internalization Ratio				

Note.—DX = depth of intrapersonal exploration.

* $p < .05$, one-tailed test.

** $p < .05$, two-tailed test.

*** $p < .01$, two-tailed test.

is able to vicariously benefit from the explorations of fellow patients within the group. In this sense, then, the level of exploration manifested by the group as a whole may somewhat accurately reflect the benefit gained by each member individually. It may be that levels of overt exploration on the part of one person leads to greater exploration on the part of the second, even though the second never voices his explorations publicly. Self-exploration, as the name suggests, may be a kind of working through of one's problems and it is in this manner that it may lead to constructive personality change.

A very interesting, but unpredicted finding has to do with the self-concept measures. Tables 1, 2, and 3 show that the therapists-within-cells terms are significant for all five *Q*-sort measures. (Actually, this term might better be called "therapy-groups-within-cells" for purposes of interpretation, rather than "therapists-within-cells," which is a handier designation.) It appears that there were differences from group to group with regard to the amount of improvement made in self-concept which is not attributable to the variables included in the study. This very specific effect, apparently independent of other personality and behavioral changes, obviously resulted from something occurring in the group process. What this might be is not clear. One might suggest any number of hypotheses, however. A likely explanation has to do with differences between therapists in their reinforcement of self-references. That is, therapists could conceivably condition positive changes in self-concepts through differential reinforcement of self-reference statements. This is likely, since self-reference statements are one indication of degree of self-exploration. This explanation receives some support from the results, presented in Table 3, which show that of the five *Q*-sort measures used, the three which are measures of self-concept against a standard all changed more in high-exploration groups, and two of these were significant. Thus, it would appear that change in self-concept is at least partially related to level of self-exploration in the therapy group.

The present study provides moderate support for the use of two procedures which differ from standard group therapy practice. Further, it suggests some relationship between self-exploration and outcome in group psychotherapy.

REFERENCES

- BLAU, B. A. A comparison of more improved with less improved clients treated by client-centered methods. In W. U. Snyder (Ed.), *Group report of a program of research in psychotherapy*. State College: Pennsylvania State College, Psychotherapy Research Group, 1953.
- BRAATEN, L. J. The movement from non-self to self in client-centered psychotherapy. Unpublished doctoral dissertation, University of Chicago, 1958.
- BUTLER, J. M., & HAIGH, G. V. Changes in the relation between self-concepts and ideal-concepts consequent upon client-centered counseling. In C. R. Rogers & R. F. Dymond (Eds.), *Psychotherapy and personality change*. Chicago: University of Chicago Press, 1954.
- FINNEY, B. C. A scale to measure interpersonal relationships in group psychotherapy. *Group Psychotherapy*, 1954, 7, 52-66.
- FORSYTHE, R. P., & FAIRWEATHER, G. W. Psychotherapeutic and other hospital treatment criteria: The dilemma. *Journal of Abnormal and Social Psychology*, 1961, 62, 598-605.
- LEWIS, M. K. Counselor prediction and projection in client-centered psychotherapy. Unpublished doctoral dissertation, University of Chicago, 1959.
- PERES, H. An investigation of non-directive group therapy. *Journal of Consulting Psychology*, 1947, 11, 159-172.
- SEEMAN, J. A study of the process of non-directive therapy. *Journal of Consulting Psychology*, 1949, 13, 157-168.
- SLAVSON, S. R. *Textbook in analytic group psychotherapy*. New York: International Universities Press, 1963.
- TOMLINSON, T. M., & HART, J. T. A validation study of the process scale. *Journal of Consulting Psychology*, 1962, 26, 74-78.
- TRUAX, C. B. The process of group psychotherapy: Relationships between hypothesized therapeutic conditions and intrapersonal exploration. *Psychological Monographs*, 1961, 75(7, Whole No. 511).
- TRUAX, C. B. Client-centered group psychotherapy. Workshop at American Group Psychotherapy Association, New York, January 1962. (a)
- TRUAX, C. B. The therapeutic process in group psychotherapy: A research investigation. Wisconsin Psychiatric Institute, University of Wisconsin, January 1962. (Mimeo) (b)
- TRUAX, C. B. A tentative scale for the measurement of depth of intrapersonal exploration (DX). In, *Discussion papers*. Wisconsin Psychiatric Institute, Madison: University of Wisconsin Press, 1962. (c)

- TRUAX, C. B., & CARKHUFF, R. R. For better or for worse: The process of psychotherapeutic personality change. In, *Recent advances in the study of behavior change*. Montreal, Canada: McGill University Press, 1963.
- TRUAX, C. B., & CARKHUFF, R. R. Personality change in hospitalized mental patients during group psychotherapy as a function of alternate sessions and vicarious therapy pretraining. *Journal of Clinical Psychology*, 1965, 21, 225-228.
- TRUAX, C. B., & WARGO, D. G. Antecedents to outcome in group counseling with institutionalized juvenile delinquents: Effects of therapeutic conditions, patient self-exploration, alternate sessions, and vicarious therapy pretraining. Unpublished manuscript, Arkansas Rehabilitation Research and Training Center, 1968. (a)
- TRUAX, C. B., & WARGO, D. G. Effects of therapeutic conditions, alternate sessions, vicarious therapy pretraining, and patient self-exploration on hospitalized mental patients during group therapy. Unpublished manuscript, Arkansas Rehabilitation Research and Training Center, 1968. (b)
- WAGSTAFF, A. K., RICE, L. N., & BUTLER, J. M. Factors of client verbal participation in therapy. In, *Counseling center discussion papers*. Chicago: University of Chicago Press, 1960.
- WOLF, A. Group psychotherapy with adults: The alternate meeting. Paper presented at the meeting of the American Personnel and Guidance Association, New York, January 1961.
- WOLF, A., & SCHWARTZ, E. K. *Psychoanalysis in groups*. New York: Grune & Stratton, 1962.

(Received August 5, 1968)

MEASUREMENT OF SOCIAL-EVALUATIVE ANXIETY¹

DAVID WATSON²

University of Hawaii

AND

RONALD FRIEND

University of Toronto

The specification of a construct of social anxiety, the subsequent development of two scales, and validation studies are reported. The two scales are the Fear of Negative Evaluation scale (FNE) and the Social Avoidance and Distress scale (SAD). The scales had very high indexes of homogeneity and possessed sufficient reliability. Three experiments and other correlational data are presented. People high in SAD tended to avoid social interactions, preferred to work alone, reported that they talked less, were more worried and less confident about social relationships, but were more likely to appear for appointments. Those high in FNE tended to become nervous in evaluative situations, and worked hard either to avoid disapproval or gain approval. Certain convergent and discriminant relationships had been part of the construct of social anxiety, and the correlational data supported these differentiations.

Several lines of research and theory point to the potential heuristic value of the development of a scale or scales to measure anxiety experienced in social situations. Those studies on the effects of social censure on the performance of psychotics (Rodnick & Garnezy, 1957), on the effect of disturbed family and social relationships in psychopathology (Gerard & Siegel, 1950), and on the correlation between social incompetence and hospitalization (Zigler & Phillips, 1962) have all suggested that distressful social relationships may be one powerful determinant of psychopathology. A second line of research has followed the idea that some individuals are more anxious than others in social situations, and might therefore be more persuasible (Sears, 1967), more concerned with others' evaluations of themselves (Diggory, 1966), or simply be upset at having to interact with other people (Byrne, McDonald, & Mikawa, 1963). A third line of evidence which points to the need for social anxiety scales is that which suggests that anxiety scales which confine their questions to specific situations or conditions, (Endler & Hunt, 1966) have greater predictive validity for those specific situations than those scales which sparsely

sample diverse situations. There do exist two sets of items which have been nominated as measuring social anxiety (Dixon, deMonchaux, & Sandler, 1957; Sears, 1967) but these items either are too few or have no underlying construct, nor have they been controlled for response style problems such as acquiescence or social desirability, and they have never been validated.

Social-evaluative anxiety was initially defined as the experience of distress, discomfort, fear, anxiety, etc., in social situations; as the deliberate avoidance of social situations; and finally as a fear of receiving negative evaluations from others. The first two aspects were combined to compose a Social Avoidance and Distress (SAD) scale, and the last factor was employed to compose a Fear of Negative Evaluation (FNE) scale.

Specification of the Constructs

The two scales were constructed at the same time. Three general goals were adopted: to suppress response style errors, to foster scale homogeneity, and to foster discriminant or convergent relationships with certain other scales. These goals can be approached by paying careful attention to the nature of the constructs from which the items follow.

Anxiety often is inferred from verbalized subjective distress, the execution of avoidance responses, impaired performance, or certain physiological signs. Almost all anxiety scales are a haphazard mixture of items asking S to

¹ This research was conducted while the first author was at the University of Toronto and was supported by Grant No. 67 from the Ontario Mental Health Foundation.

² Requests for reprints should be sent to David Watson, Psychology Department, University of Hawaii, Honolulu, Hawaii 96822.

report on these four aspects of his behavior or experience. It was decided in the construction of the present scales to exclude items asking about physiological signs or impaired performance. This would make clear what behavior was used as a sign for membership in the class "anxious"—subjective distress and avoidance—and what behavior was considered a correlate of class membership—impaired performance and physiological signs. The advantage of this specification is that one may later more easily distinguish between construct and predictive validity.

A second, rarely followed, requisite for careful delineation of a construct is the necessity for adequate definition of the range of the trait, particularly its opposite instance (Jackson, 1966a). To control for acquiescence response set in a true-false format, as was done here, one must have approximately half of the items worded so that answering "false" indicates presence of the trait, which requires an adequate definition of the opposite instance. For example, the opposite instance of social avoidance is not necessarily social approach, that is, affiliation, but simply a lack of avoidance. If the opposite instance of avoidance were allowed to be affiliation it would have obscured any possible discriminant relationship between the SAD scale and measures of affiliation. Particular attention was paid to wording items so that the opposite instance of a trait simply indicated absence of that trait, not the presence of some other trait.

Fear of negative evaluation was defined as apprehension about others' evaluations, distress over their negative evaluations, avoidance of evaluative situations, and the expectation that others would evaluate oneself negatively. The presence of high FNE does not necessarily imply that an individual evaluates himself negatively, or that he is concerned about revealing his inferiority (Dixon et al., 1957). Fear of loss of social approval would be identical to FNE, but the opposite instance of FNE is simply lack of anxiety about others' evaluations, not inevitably a desire for positive evaluation. High FNE differs from test anxiety in that it is not specific to testing conditions, but may operate in any

social, evaluative situation, such as being on a date, talking to one's superiors, or being interviewed for a job.

The SAD scale was divided into two subscales, social avoidance and social distress. The purpose was to create a general scale, so the respondent was not asked *why* he experienced distress or avoided social encounters. Social avoidance was defined as avoiding being with, talking to, or escaping from others for any reason. Both actual avoidance and the desire for avoidance were included. The opposite instance was simple lack of an avoidance motive, not desire to affiliate. Social distress was defined as the reported experience of a negative emotion, such as being upset, distressed, tense, or anxious, in social interactions, or the reported lack of negative emotion, such as being relaxed, calm, at ease, or comfortable. The opposite instance of distress was lack of unhappiness, not the presence of some positive emotion.

The item selection procedure was explicitly chosen to eliminate as much social desirability variance as possible, for questions about anxiety in social situations would tend to elicit great concern with issues of social desirability, and the additional variance would not have been helpful. In two other instances discriminant relationships were explicitly encouraged. Questions about S's reaction to testing situations were excluded, as there are already several test-anxiety scales in existence, and because the social anxiety scales could be more explicitly oriented to social situations. Finally, the authors attempted to foster a discriminant relationship between the SAD and FNE scales themselves, in the hope that separating them would increase their heuristic value.

CONSTRUCTION OF THE SCALES

Following the stringent criteria for evaluation of items suggested by Jackson (1966a), 145 items were selected by rational analysis from a much larger pool.³ These were then subjected to an empirical test, and the final 58 items which survived as the two scales

³ The authors wish to thank Barney Gilmore, Stewart Page, and Sybil Paterson, who either contributed suggestions for items or aided in their evaluation.

are presented in Tables 1 and 2. The scoring key is given after each item. The SAD scale is evenly divided between true and false items; the FNE scale has 17 true and 13 false items.

The 145 items initially selected were administered to 297 undergraduates at the Uni-

TABLE 1

SOCIAL AVOIDANCE AND DISTRESS (SAD) SCALE

1. I feel relaxed even in unfamiliar social situations. (F)
2. I try to avoid situations which force me to be very sociable. (T)
3. It is easy for me to relax when I am with strangers. (F)
4. I have no particular desire to avoid people. (F)
5. I often find social occasions upsetting. (T)
6. I usually feel calm and comfortable at social occasions. (F)
7. I am usually at ease when talking to someone of the opposite sex. (F)
8. I try to avoid talking to people unless I know them well. (T)
9. If the chance comes to meet new people, I often take it. (F)
10. I often feel nervous or tense in casual get-togethers in which both sexes are present. (T)
11. I am usually nervous with people unless I know them well. (T)
12. I usually feel relaxed when I am with a group of people. (F)
13. I often want to get away from people. (T)
14. I usually feel uncomfortable when I am in a group of people I don't know. (T)
15. I usually feel relaxed when I meet someone for the first time. (F)
16. Being introduced to people makes me tense and nervous. (T)
17. Even though a room is full of strangers, I may enter it anyway. (F)
18. I would avoid walking up and joining a large group of people. (T)
19. When my superiors want to talk with me, I talk willingly. (T)
20. I often feel on edge when I am with a group of people. (T)
21. I tend to withdraw from people. (T)
22. I don't mind talking to people at parties or social gatherings. (F)
23. I am seldom at ease in a large group of people. (T)
24. I often think up excuses in order to avoid social engagements. (T)
25. I sometimes take the responsibility for introducing people to each other. (F)
26. I try to avoid formal social occasions. (T)
27. I usually go to whatever social engagements I have. (F)
28. I find it easy to relax with other people. (F)

TABLE 2

FEAR OF NEGATIVE EVALUATION (FNE)

1. I rarely worry about seeming foolish to others. (F)
2. I worry about what people will think of me even when I know it doesn't make any difference. (T)
3. I become tense and jittery if I know someone is sizing me up. (T)
4. I am unconcerned even if I know people are forming an unfavorable impression of me. (F)
5. I feel very upset when I commit some social error. (T)
6. The opinions that important people have of me cause me little concern. (F)
7. I am often afraid that I may look ridiculous or make a fool of myself. (T)
8. I react very little when other people disapprove of me. (F)
9. I am frequently afraid of other people noticing my shortcomings. (T)
10. The disapproval of others would have little effect on me. (F)
11. If someone is evaluating me I tend to expect the worst. (T)
12. I rarely worry about what kind of impression I am making on someone. (F)
13. I am afraid that others will not approve of me. (T)
14. I am afraid that people will find fault with me. (T)
15. Other people's opinions of me do not bother me. (F)
16. I am not necessarily upset if I do not please someone. (F)
17. When I am talking to someone, I worry about what they may be thinking about me. (T)
18. I feel that you can't help making social errors sometimes, so why worry about it. (F)
19. I am usually worried about what kind of impression I make. (T)
20. I worry a lot about what my superiors think of me. (T)
21. If I know someone is judging me, it has little effect on me. (F)
22. I worry that others will think I am not worthwhile. (T)
23. I worry very little about what others may think of me. (F)
24. Sometimes I think I am too concerned with what other people think of me. (T)
25. I often worry that I will say or do the wrong things. (T)
26. I am often indifferent to the opinions others have of me. (F)
27. I am usually confident that others will have a favorable impression of me. (F)
28. I often worry that people who are important to me won't think very much of me. (T)
29. I brood about the opinions my friends have about me. (T)
30. I become tense and jittery if I know I am being judged by my superiors. (T)

versity of Toronto. The Crowne-Marlowe (1964) Social Desirability scale and the first 10 items of Jackson's (1966b) Infrequency scale were also administered. The latter was used to control for pseudorandom responding, and Ss who answered any of these items incorrectly were dropped from all analyses. Ninety-two Ss were dropped for this reason or because they did not answer all items.

The Crowne-Marlowe scale provided an initial empirical criterion against which to evaluate the SAD and FNE items. Jackson's (1966a, 1967) item selection procedure was employed, to minimize covariation with social desirability as a response style. For every item for 205 Ss a computer calculated the biserial correlation of the item with its own scale and the biserial correlation with all other scales in the sample. Items endorsed by less than 5% of the sample were dropped. Jackson's Differential Reliability Index (DRI) was calculated for each remaining item. In the first instance, the DRI was calculated for the item's own scale and the Crowne-Marlowe, giving an estimate of the amount of correlation between an item and its own scale with social desirability variance removed. The first criterion for selecting items was that this DRI be as high as possible and in all cases above .50. It would not have been possible to apply further criteria had not the original rational analysis provided a large number of items which passed this first hurdle.

The second criterion was that the probability of endorsement of each item must have been above 10% and as close to 50% as possible. Third, a discriminant relationship between the SAD and FNE was encouraged by using a DRI in which item-own scale correlation was estimated with variance attributable to the other scale removed, selecting those items with minimal common variance. Fourth, to control for acquiescence, items were selected by judging the adequacy of those representing the opposite instance of the trait. Finally, items were selected according to content dissimilarity and the representativeness of the situations about which the items inquired.

Characteristics of the Scales

One of the major goals was to foster scale homogeneity. This can be expressed as the mean biserial correlation of each item with its own scale. The mean biserial correlation of the selected FNE items, corrected for presence of item in the total score, was .72; and of the selected SAD items, it was .77 ($N = 205$, $p < .01$). The product-moment correlation of the two subscales of the SAD, avoidance and distress, was .75. A second index of homogeneity is the KR-20 reliability statistic. The KR-20s of both the selected FNE and SAD items were .94. The KR-20 of the Crowne-Marlowe was .79. In a second sample of 154 Ss the KR-20 of the FNE was .96 and that of the SAD was again .94. The two scales are very homogeneous.

A second goal was to minimize the relationship of the scales to social desirability. The product-moment correlations of both the FNE and SAD with the Crowne-Marlowe scale were $-.25$ ($N = 205$, $p < .01$). These correlations might have been less had the Crowne-Marlowe itself been more homogeneous. In any case, the relationship with social desirability has been minimized. Another goal had been to foster a discriminant relationship between FNE and SAD. In the first sample, the product-moment correlation between the two was .51 ($N = 205$, $p < .01$). A later sample of 42, with which a Spearman correlation was used, showed a correlation of .32 ($N = 42$, $p < .05$). The attempt to foster a discriminant relationship between SAD and FNE did not eliminate all the common variance, probably because some people score highly on SAD just because they are fearful in social-evaluative situations. However, the common variance has been minimized.

Neither of the scales was normally distributed. The distribution of the FNE was nearly rectangular. The mean was 15.47, the standard deviation was 8.62, and the median was 16. This rectangularity indicated that the entire range of the scale was well used. The distribution of the SAD was skewed. Although the modal score was zero, the mean was 9.11, the median was 7, and the standard deviation was 8.01. Separate analysis of the social avoidance and social distress subscales

indicated that the distribution of social avoidance was most skewed. High levels of SAD were not as common as high levels of FNE. High levels of SAD may be more pathological. Variables determining extreme social withdrawal or distress, which might be termed schizoid, are probably not normally distributed within the general population, which would explain the skew in the SAD data. This lack of normality in the SAD may make easier the task of early identification of those who later show schizoid reactions.

There were differences between the sexes in scores on the two scales. The mean scores on the SAD were: males, ($N = 60$) 11.20; females ($N = 145$) 8.24. This difference is significant ($t = 2.64$, $p < .01$). Males reported more social avoidance and distress than females. The mean scores on the FNE were: males, 13.97; females, 16.10. This difference approached significance (N the same as before; $t = 1.76$, $p < .10$). In this case the direction was reversed, women reporting more fear of negative evaluation than men.

Data for test-retest reliability were gathered on a sample of 154 Ss in the summer school of the University of Toronto. The SAD and FNE scales were administered alone, during class time, and 1 month passed between administrations. The product-moment, test-retest correlation of the FNE was .78, and that of the SAD was .68. A second sample of 29 gave the figures .94 for FNE and .79 for SAD, but the larger size of the first sample makes it a better estimate. These figures indicated sufficient reliability.

EXPERIMENTAL STUDIES⁴

The SAD and FNE scales were administered to 358 students in the summer school at the University of Toronto. From this pool Ss were drawn for three experiments, whose purpose was to validate the two scales.

SAD and the Prospect of Working Alone or Together

Persons who are high on SAD would be expected to be uncomfortable in social situa-

tions, and to prefer to be alone. Even the prospect of having to be in a future social interaction might make those high on SAD anxious. In this study Ss were told that they would later be required to participate in a group discussion or write an essay alone, and the effect of this expectation on their performance, anxiety, and other attitudes was observed.

From the 82 Ss scoring zero or one on the SAD, 46 were randomly selected as the low anxious (LA) group, and from the 85 scoring 12 or above, 52 were selected as the high anxious (HA) group. Within these two groups Ss were randomly assigned either to an Essay or Group Discussion condition. The Ss participated in the study in heterogeneous groups varying in size from 6 to 15. All Ss had the same female *E*. The Ss were told that they would actually be in two experiments, one to occur immediately and the second afterward. In the Essay condition, Ss were informed that in the second experiment they would be placed alone in a cubicle where they would write an essay on an interesting, controversial topic. In the Group Discussion condition, Ss were told that they would be placed in a small group in which they would be expected to participate actively in a discussion of an interesting, controversial topic. Envelopes labeled "Essay Writing" or "Group Discussion" which ostensibly contained material for the later work were provided. The Ss then participated in what was for them the first of two experiments, in which they had administered to them the Digit Symbol subtest of the Wechsler Adult Intelligence Scale. If the Ss were very anxious, this would have been reflected in impaired performance on this test. The Ss then filled out a questionnaire. After this *E* said that the second experiment could not take place at that time, and elicited an indication of how interested Ss might be in returning for the second study. That concluded the experiment. All of the significant results came from the questionnaire, as the simple expectation of later social interaction apparently did not arouse the degree of anxiety necessary to interfere with performance at the Digit Symbol task.

The Ss indicated how interested they were in returning at a later date by checking a 5-point scale labeled "Not at all" at one end,

⁴ The authors wish to thank Betty Jean Findlay, Donna Hutchinson, and Mary Ann Linseman for serving as *Es*. The authors appreciated the useful advice of John Arrowood on the design of one of the studies.

TABLE 3
QUESTIONNAIRE DATA FROM EXPERIMENT I

Item	Condition	
	Essay	Group Discussion
Degree of interest in returning		
HA	2.67	2.39
M	24	28
N		
LA	3.19	3.44
M	21	25
N		
Number of Ss choosing each condition		
HA*	27	24
LA	5	41
Degree of concern		
HA	2.33	1.89
LA	1.28	1.16
Degree of confidence		
HA	3.20	2.64
LA	3.28	3.68

* One S did not respond.

and "Very much" at the other. Analysis of variance for unequal N was performed on these and all other data. Table 3 presents the data. The difference between the HA and LA groups ($F = 9.49$, $df = 1/94$, $p < .01$), indicated that as expected the HA Ss were less interested in returning. The overall difference between the Essay and Group Discussion conditions was also significant, but more interesting was the interaction ($F = 45.73$, $df = 1/94$, $p < .001$), which indicated that while HA Ss were more interested in returning if they were to be in the Essay condition, the LA Ss were more interested if they were to be in the Group Discussion. The Ss were also asked to indicate their choice of being in the Essay or Group Discussion conditions when they returned. The data, also in Table 3, were analyzed by chi-square ($\chi^2 = 19.11$, $df = 1$, $p < .001$). The LA Ss showed a much greater preference for the Group Discussion condition. These two analyses pointed up much greater interest in being sociable among LA Ss.

The Ss were asked how worried or concerned they were about the second experiment, and they responded on a 5-point scale as above. The HA Ss were more worried ($F = 26.38$, $df = 1/94$, $p < .001$). The Essay

writing condition produced more concern than the Group Discussion condition ($F = 7.39$, $df = 1/94$, $p < .01$), perhaps because an essay is tangible and more easily evaluated while one's participation in a group discussion is not. If this is so, then it would explain the interaction effect ($F = 4.25$, $df = 1/94$, $p < .05$), in the direction opposite that expected, as the HA Ss were most concerned in the Essay condition. The HA Ss in that condition may have been worried about some possible later evaluation of their essay.

Another question asked Ss to predict how uneasy or nervous they expected to feel in the second experiment, and they responded on a 5-point scale. There was a difference between the anxiety groups ($F = 26.77$, $df = 1/94$, $p < .001$), but no other effects. The means were: HA, 3.02 and LA, 1.80. In another question using a 5-point scale, LA Ss reported that they expected to be calmer than HA Ss did in the coming experiment ($F = 14.46$, $df = 1/94$, $p < .001$). The means were: HA, 2.98 and LA, 3.80. The HA Ss expected to be more nervous, less calm, in the second experiment. A final question asked about the degree of apprehension felt about participating in the coming study. A 5-point scale was labeled "apprehensive" and "confident" at its extremes. The means are in Table 3. The LA Ss reported more confidence ($F = 6.70$, $df = 1/94$, $p < .05$), and there was an interaction ($F = 4.92$, $df = 1/94$, $p < .05$) which showed the expected pattern: HA Ss were more confident or less apprehensive if they were to be in the Essay condition, while LA Ss were more confident if assigned to the Group Discussion.

To sum up, people who scored highly on the SAD scale were less likely to be interested in returning, and more likely to choose to be alone, both indexes of social avoidance. The HA Ss were more worried and uneasy as well as less calm and confident—all indexes of distress—and these discomforts were magnified in the together condition. Two other studies provided modest, additional validation for the two scales.

SAD and Being Alone or Together

From the same distribution of SAD scores, 97 HA Ss, with scores of 11 and above, and 58 LA Ss, with scores below three, all of whom

performed at the same tasks, were randomly assigned to work either alone, together in a cubicle with another person but without talking, or together and talking to the other person in the cubicle. The tasks were anagrams, risk-level, and the learning of paired-associative nonsense syllables. The experimental conditions produced results in the predicted directions, but were too weak to elicit enough anxiety to show significant effects on performance. In the Talking condition, the *Ss* were encouraged to talk, but how much they actually talked was left up to them. When asked to indicate how much they had actually talked on a scale, labeled "Briefly, occasionally, frequently, and almost all the time" at its four points, HA *Ss* reported significantly less talking ($t = 2.20$, $df = 42$, $p < .05$). The means were: HA, 2.7; LA, 3.3. Not talking to others is a kind of social avoidance, and the SAD scale is supported here in that it differentiates *Ss* according to how much they report talking to others.

FNE, Approval Seeking or Disapproval Avoiding

If fear of negative evaluation is an avoidance motive, then a person high in FNE might try to gain social approval simply as a way of avoiding disapproval. Individuals high on FNE might be expected to be most affected by the possibility of disapproval, while those low on FNE might be more affected by the possibility of gaining approval. This study attempted to manipulate these incentive conditions.

The *S*, seated alone in a cubicle, believed that the work he was doing was part of a group effort, and that his performance would be evaluated by an unseen, unknown group leader. From the top 25% of the same FNE distribution—scores above 18—and from the bottom 25%—scores less than 9—48 high FNE and 48 low FNE *Ss* were randomly selected. Within each of these two groups *Ss* were randomly assigned either to a condition in which they were told that their group leader might approve of their work, but would never disapprove of it (Approval), or to a condition in which the leader might disapprove, but would never approve (Disapproval). Actually there were no groups and

Ss never heard from their leader. The *Ss* performed a letter-number substitution task for 4 minutes, the dependent variable being the number of items completed. Analysis of variance indicated that the only effect which approached significance was the interaction pattern ($F = 3.76$, $df = 1/92$, $p < .10$). The mean scores were: high FNE Approval, 148.2; high FNE Disapproval, 145.2; low FNE Approval, 132.9; low FNE Disapproval, 149.5. If this unexpected, borderline effect may be believed, it indicates that those high on FNE not only avoided disapproval but sought approval, while the low FNE *Ss* were not approval seekers, and were motivated to work hard on a relatively dull task only if disapproval were threatened.

On a questionnaire, *Ss* were asked how nervous or uneasy they felt during the experiment, and they responded on a 5-point scale. As expected, those high on FNE reported more uneasiness ($F = 17.01$, $df = 1/92$, $p < .001$). The means were: high FNE, 2.5; low FNE, 1.6. The *Ss* had also been asked how they evaluated their own performance and how they thought their group leader would evaluate their performance. Neither of these questions revealed differences between groups, but some *Ss*—18 out of 96—reported that they expected the group leader to think less well of their work than they did. Of the 18, 10 were high FNE and 8 were low, which is a chance occurrence, but all 10 of the high FNE *Ss* were in the Approval condition, while the low FNE *Ss* were evenly divided between Approval and Disapproval. By Fisher's exact test this pattern is significantly different from chance ($p < .025$). If this post hoc analysis indicates a real difference, then it suggests a defensive sort of expecting the worst among some high FNE *Ss*.

CORRELATIONAL STUDIES

To check on certain discriminant or convergent relationships which had been specified, and to search for new relationships, the SAD and FNE scales were correlated with other measures. Table 4 presents the correlations of FNE and SAD with Taylor's (1953) Manifest Anxiety (*MA*) scale, Rotter's

TABLE 4

CORRELATION OF FNE AND SAD WITH OTHER SCALES

Original sample*				
	MA	LC	AH — ^a	E-H**
FNE	.60	.18	.28	.47
N	171	172	173	35
SAD	.54	ns	.18	.45
N	171		173	35

Second sample**							
	ASI	Social Ap- proval	Affil- iation	Desir- ability	Auton- omy	Defend- ence	Achieve- ment
FNE	.39	.77	ns	-.58	-.32	.42	ns
SAD	.76	ns	-.76	-.42	ns	ns	-.33

Third sample**					
	Aggres- sion	Domi- nance	Abase- ment	Exhibi- tionism	Impul- sivity
FNE	ns	-.50	.29	-.39	ns
SAD	ns	ns	ns	ns	ns

Note.—For the second sample, $N = 42$; for the third, $N = 40$.

* There were no significant correlations with AH+.

** If $r = .15$, $p < .05$; if $r = .19$, $p < .01$.

** If $r = .30$, $p < .05$; if $r = .39$, $p < .01$.

(1966) Locus of Control scale (LC), Alpert and Haber's (1960) Achievement Anxiety scale, divided into debilitating and facilitating anxiety subscales (AH— and AH+), the social and evaluative parts of the Endler-Hunt (1966) S-R Inventory of Anxiousness (E-H), Paivio's (1965) Audience Sensitivity Index (ASI), and 11 of the subscales of Jackson's (1966a, 1966b) Personality Research Form (Social Approval, Affiliation, Desirability, Autonomy, Defence, Achievement, Aggression, Dominance, Abasement, Exhibitionism, and Impulsivity). The table presents only significant or borderline correlations.

There should be some common variance between a general and a more specific measure of anxiety. If the MA scale is a measure of general anxiety, then the FNE and SAD should have a moderate relationship to it, which was the case. An attempt was made to foster a discriminant relationship with test anxiety, and the low correlation of both the social anxiety scales with AH— indicates that this was accomplished. If the situation eliciting the state of anxiety is as important or even more important than the trait itself (Endler & Hunt, 1966), then fostering dis-

criminant relationships between situationally specific scales is valuable. Indeed, Cherlin (1967) showed that, compared with other anxiety measures, only a measure of state anxiety in social situations correlated with behavior in those situations.

A scale measuring social avoidance would have to be negatively correlated with affiliation, and the SAD met this criterion. Whether lack of affiliation necessarily implies presence of social anxiety is a moot point. The problem is that a person low on both traits would answer the opposite instance items of the two scales similarly, thus producing a spurious correlation, that is, a correlation between scales which did not imply a relationship between constructs. The present data indicate that it is impossible to be high on both traits, which is as it should be, given the nature of the constructs, but whether it is possible to be low on both traits cannot definitely be answered by these data. This possible spurious correlation would be a general problem in specifying the relationship between scales which control for acquiescence by employing opposite instance items which inquire simply about the absence of the trait in question. Yet the paradox is that the best kind of item does just that to avoid confounding with another trait.

The Endler-Hunt items which were used here ask *S* to report how he would respond in social-evaluative situations such as going to meet a new date, going on an interview for a very important job, going into a psychological experiment, etc. The moderate correlation of the SAD and FNE scales with reported reactions in these situations supports the validity of the two scales.

Paivio's ASI scale inquires about both the reaction of *S* to audience situations and possible causes for that reaction. The evaluative nature of audience situations is indicated by the correlation of FNE and ASI. Note also the negative relationship of FNE and Exhibitionism. People who fear others' evaluations don't want to be the center of attention. The very high relationship between SAD and ASI delineates the experience of stage fright as intense discomfort and a desire to flee.

The high correlation of FNE and Social

Approval was consistent with the tentative results of the third experiment, suggesting that people high in FNE are very concerned with gaining others' approval. It should be noted that one item is identical on the two scales, and that the wording of the opposite instance items again raises the possibility of a spuriously high correlation. High FNE may imply a desire for social approval, but it is not clear that the latter necessarily implies FNE, nor is it known if it is possible to be low on both traits. This correlation does not imply that FNE is heavily loaded with social desirability as a response set, which was minimized by the use of the Marlowe-Crowne scale in the item selection procedure. Rather, people who fear other's negative evaluations want their approval as a signal that the feared outcome has been avoided.

Jackson's Desirability scale, on its face, seems to measure the degree to which a person actually engages in socially undesirable behavior. Presumably both high levels of SAD and FNE are somewhat undesirable, hence the moderate relationship with Desirability. Note that this is not desirability as a response style, which was assessed by the Crowne-Marlowe scale. The remainder of the significant correlations gave validation support to the FNE and SAD scales in a fairly straightforward manner. People high on FNE tend to be defensive, not autonomous, not dominant, and perhaps self-effacing. Some of the Defence items, in fact, inquire about defensiveness in evaluative situations. To be autonomous or dominant opens one to criticism which is exactly what the person high on FNE fears. The negative correlation of SAD and Achievement emphasizes the social nature of the competitive achievement motive.

One other validation study will be briefly reported. Summer school students at the University of Toronto were notorious for not keeping appointments to serve as Ss in experiments. A test was made to see if this behavior was related to social anxiety. Fifty-two of the 81 high SAD Ss, with scores above 10, versus 36 of 80 low SAD Ss, scores below 3, appeared for an experimental appointment ($\chi^2 = 5.99$, $df = 1$, $p < .02$). There were no effects of FNE. Apparently these high SAD

Ss were sufficiently afraid of the E or their professor (neither of whom had any power to enforce appearance for appointments) or the generalized other, to do what was asked of them. People high in SAD may be more ingratiating or conforming than others.

DISCUSSION

As expected, people high on SAD did avoid social situations, and were anxious in social interactions. Individuals high on FNE became nervous in evaluative conditions, and seemed to seek social approval. Both scales showed correlations with other relevant measures. The results lend validity to the two scales.

It is worth mentioning constructs which might have been embodied in the two scales, but were not: fear of revealing one's inferiority, concern about the appearance of one's body, fear of lack of control of bodily processes, fear of one's unmanageable motives, inadequate expectations concerning what others will do; not knowing how to behave in social situations. These and other notions may express some of the origins of social anxiety. They were not included in the present construct to avoid begging the question of origins, but now that relatively unambiguous measures of the trait have been developed it will be possible to explore the relationships between variables. For example, one may ask how fear of other's evaluations is related to self-evaluation.

Social anxiety is a variable which may act as a moderator in a number of social situations. For example, the threat of negative evaluation might increase the chances of eliciting compliant behavior if the individual is in a state of anxiety and appraises that he can reduce anxiety through compliance. Thus FNE or SAD might act as moderator variables in studies on ingratiation, conformity, persuasibility, the demand characteristics of experiments, social approval seeking, social comparison processes, etc. Studies of the social processes within groups, such as the study of communication, might use the SAD as a moderator variable. In general, the FNE might be relevant in any potential evaluative situation, and the SAD could be useful in studies of social interaction.

The situation eliciting social anxiety is important, and one would expect to find that different situations have differential effects on individuals. The potential item universe from which items were selected for the FNE and SAD was intended to be all social situations, though only a few situations were unsystematically sampled by the items actually chosen. The sex of the person with whom *S* is interacting, his status relative to *S*, whether he is a friend or stranger, in public or private conditions, under evaluative or nonevaluative circumstances, with covert or overt evaluation—all of these varied conditions will differentially affect people high or low in social anxiety. In an evaluative situation, how do the characteristics of the evaluator affect a person's anxiety level, and how does the behavior of the anxious person affect the judgment of the evaluator? Social avoidance and distress, measured by the SAD, may be a general reaction for some people, while for others it may be specific to certain conditions, such as dealing with authorities or members of the opposite sex.

Very little is known about the development of high levels of social anxiety, or about what affect they have on individuals. The amount of FNE may be a simple function of the amount of prior disapproval received, or a more complicated function combining prior disapproval with desire for approval. It seems best to conceive it as acquired on the basis of some sort of frustration or punishment. The relationship between approval seeking and disapproval avoiding is unclear. Individuals with high levels of FNE may seek nonevaluative social situations. They may fear not only obvious evaluative conditions, but also expect and fear evaluation where none is intended. The amount of SAD may also be a function of prior frustration or punishment in social interaction. The motives or frustrations underlying the acquisition of high SAD are unclear. A person with high SAD is clearly isolated and often fearful. The relationship of social anxiety to psychopathology is hinted here: the SAD scale may make it possible to study not simply the schizoid versus the normal, but intermediate types as well. Hopefully the development of the FNE and SAD scales will

make it possible to study several of the interactions of personality and social variables discussed here.

REFERENCES

- ALPERT, R., & HABER, R. N. Anxiety in academic achievement situations. *Journal of Abnormal and Social Psychology*, 1960, 61, 207-215.
- BYRNE, D., McDONALD, R., & MIKAWA, J. Approach and avoidance affiliation motives. *Journal of Personality*, 1963, 31, 21-37.
- CHERLIN, D. L. Anxiety and consultant differences in self-study groups. Unpublished doctoral dissertation, Yale University, 1967.
- CROWNE, D. P., & MARLOWE, D. *The approval motive: Studies in evaluative dependence*. New York: Wiley, 1964.
- DIGGORY, J. C. *Self-evaluation, concepts and studies*. New York: Wiley, 1966.
- DIXON, J. J., DEMONCHAU, C., & SANDLER, J. Patterns of anxiety: An analysis of social anxieties. *British Journal of Medical Psychology*, 1957, 30, 107-112.
- ENDLER, N. S., & HUNT, J. McV. Sources of behavioral variance as measured by the S-R Inventory of Anxiousness. *Psychological Bulletin*, 1966, 65, 336-346.
- GERARD, D. L., & SIEGEL, J. The family background of schizophrenia. *Psychiatric Quarterly*, 1950, 24, 47-73.
- JACKSON, D. N. A modern strategy for personality assessment: The Personality Research Form. Research Bulletin No. 30, 1966, Psychology Department, University of Western Ontario. (a)
- JACKSON, D. N. *Personality research form*. Goshen, N. Y.: Research Psychologists Press, 1966. (b)
- JACKSON, D. N. Acquiescence response styles: Problems of identification and control. In I. A. Berg (Ed.), *Response set in personality assessment*. Chicago, Aldine, 1967.
- PAIVIO, A. Personality and audience influence. In B. Maher (Ed.), *Progress in experimental personality research*. Vol. 2. New York: Academic Press, 1965.
- RODNICK, E., & GARMEZY, N. An experimental approach to the study of motivation in schizophrenia. *Nebraska Symposium on Motivation*, 1957, 5, 109-184.
- ROTTER, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 1966, 80(1, Whole No. 609).
- SEARS, D. O. Social anxiety, opinion structure and opinion change. *Journal of Personality and Social Psychology*, 1967, 7, 142-151.
- TAYLOR, J. A personality scale of manifest anxiety. *Journal of Abnormal and Social Psychology*, 1953, 48, 285-290.
- ZIGLER, E., & PHILLIPS, L. Social competence and the process-reactive distinction in psychopathology. *Journal of Abnormal and Social Psychology*, 1962, 65, 215-222.

(Received August 12, 1968)

EFFECT OF PHYSICAL AND COGNITIVE SOMATIC AROUSAL ON RORSCHACH RESPONSES:

AN EXPERIMENTAL TEST OF THE ASSUMPTION THAT BODY IMAGE INFLUENCES THE PERCEPTUAL ORGANIZATION OF UNSTRUCTURED STIMULI

STEVEN N. VAN DE MARK¹ AND CHARLES NEURINGER²

University of Kansas

An attempt was made to experimentally verify the assumption that body image influences the perceptual organization of unstructured stimuli. It was first established that the stimulus manipulation procedures used in this study produced adequate somatic focus effects. Then male and female introductory psychology students were randomly assigned into six somatic focus arousal conditions. After being administered the somatic focus arousal stimuli, Ss were administered the Rorschach test which was scored for Barrier and Penetration Index scores. It was found that the externally aroused somatic focus Ss produced significantly higher Barrier scores than did the other Ss and that significantly more Penetration responses were made by the internal somatic focus group when compared with the other groups. It was concluded that the results of this study verified the underlying assumption that body image states influence perception.

Fisher and Cleveland (1958, 1965) feel that one's body image is crucial in the separation of self and world, is fundamental for the establishment of personal identity, and is basic in the development of adjustment strategies. They have proposed two Rorschach scoring procedures to assess body image boundary status. The basic assumption underlying their Rorschach scoring method is that one's body image influences perceptual organization of unstructured stimuli.

The body image status assessment procedure developed by Fisher and Cleveland is in terms of (a) a Barrier Index in which Rorschach responses emphasizing an external somatic locus (e.g., "rocky walls," "person covered up with a blanket," "mummy wrapped up," etc.) reflect the existence of a definite body-world boundary, and (b) a Penetration Index in which responses emphasizing an internal somatic locus (e.g., "bullet piercing flesh," "X-ray of the body," "dog eating a bone," etc.) reflect indefinite body image.

An impressive amount of research (Fisher, 1963; Fisher & Cleveland, 1965), mostly correlational, has been done using these two indexes. However, questions have arisen concerning Fisher and Cleveland's assumption that the body image influences perception of unstructured stimuli on the Rorschach test. Mednick (1959) and Wylie (1961) have argued that the Barrier and Penetration scores may reflect general modes of dealing with Rorschach inkblots rather than being the result of mediating body image experiences. Fisher and Cleveland (1965) have acknowledged these criticisms and stated that "available information does not at this time permit a meaningful answer [p. 65]." However, Fisher (1963) does feel that the underlying assumption has some validity since much of his, and of his associate's research, has established solid correlations between body phenomena and the two Rorschach indexes. Several studies (Fisher, 1965; Fisher & Fisher, 1964) have related index scores to awareness of body sensations. Unfortunately, these studies were correlational in nature and thus do not provide a critical test of the assumption that body sensations initiate particular effects to appear on the Rorschach. These researches only indicated that body sensation

¹Now at the Austin Mental Health Center, Austin, Texas.

²Requests for reprints should be sent to Charles Neuringer, Department of Psychology, University of Kansas, Lawrence, Kansas 66044.

awareness and certain kinds of Rorschach responses co-vary together. Two later attempts (Fisher & Renik, 1966; Renik & Fisher, 1968) to study the effects of concentration on the interiors and exteriors of the body found that concentrating on body interiors lowered Barrier scores in males; females who concentrated on body exteriors raised their Barrier scores.

A critical test of the assumption that the body image influences perceptual organization of unstructured stimuli would have (a) to be based on actual presentation and manipulation of arousal conditions that would lead Ss to experience their bodies in certain known ways and (b) to then relate these experimentally induced states to Rorschach responses. If it can be demonstrated that the experimental manipulation of body awareness focus produces concomitant Rorschach responses, then Fisher and Cleveland's underlying assumption is substantially corroborated. Such a demonstration would also support the wider contention (e.g., Freud, 1927; Wapner & Werner, 1965; Witkin, 1965) that the matrix of body experiences in one's life, referred to as the body image, can contribute to the factors mediating and influencing one's perception of self and world.

In this study an attempt was made to manipulate physical and cognitive arousal procedures in such a way as to induce external or internal somatic focus and to ascertain whether these affected Barrier and Penetration scores.

METHOD

In order to evaluate the effect of experimentally induced body focus states on Rorschach responses, it must first be established that the techniques used to arouse somatic focus are adequate. Without the establishment of their adequacy, it would be difficult to relate Rorschach responses to body image focus. Failure to demonstrate a relationship between body focus and Rorschach indexes could either be due to the actual nonexistence of the relationship, insensitivity on the part of the Rorschach, or to ineffective experimental manipulations of the independent variable. In order to avoid this *cul-de-sac*, a validating study of the experimental manipulations was carried out before evaluating their effect on the Rorschach.

Subjects

All Ss in the validating of the arousal procedure study and the investigation of their effect on Ror-

schach responses were chosen at random from introductory psychology classes at the University of Kansas during the 1968 spring semester. They were randomly assigned to one of six arousal conditions in such a manner as to insure equal sex representation in each group. All of the Ss went through the routines on an individual basis. The scoring of the independent variable measures was carried out without knowledge of S's group assignment. Scoring reliability evaluations were made for all the measures.

Arousal Procedures

An attempt was made to develop procedures that would physically and cognitively induce internal and external somatic focus.³ A set of procedures were also developed for the control group. These latter manipulations were designed to induce focus on environmental aspects which had no direct reference to body procedures. Half of the procedures attempted to induce arousal through physical manipulations. For each of these sets of physical procedures there was a comparable condition in which Ss were asked to imagine carrying out or experiencing the activities for equivalent amounts of time. There were six arousal conditions.

Internal focus arousal. In the physical internal focus arousal condition, S swallowed a small cup of crushed ice, sniffed from a bottle of liquid Bar-B-Q smoke for 1 minute, listened to his own heart beat for 1 minute, and drank a small cup of warm cola. In the comparable cognitive arousal condition, S was asked to imagine doing the things described above for the same amounts of time.

External focus arousal. In the physical external focus arousal situation, Ss were asked to hold arms horizontally straight outward while sitting for 1 minute, hold both hands under cold water for 1 minute, hold both legs horizontally straight outward while sitting for 1 minute. In addition, Ss also vibrated their neck, shoulders, and arms for 1 minute with an electric hand vibrator. In the comparable cognitive arousal condition, Ss were asked to imagine doing these things for identical periods of time.

³ The selection of procedures that would lead to only internal or external somatic focus is extremely difficult since these two body areas are not isolated from each other and seem to have interactive properties. Primary stimulation of one of these sectors is accompanied by secondary arousal of the other. One of the external focus arousal procedures used in this study (holding arms out straight for 1 minute), while stimulating the periphery, also gives rise to internal muscle sensations. In the same manner, a procedure which focuses on internal arousal has secondary stimulation effects at the periphery of the body. The selection of arousal procedures can only be based on the presumed *primary* locus of arousal. The presumed primary locus of arousal can be partially validated by phenomenological experience. For example, holding hands or legs out straight is reported by Ss to effect their "outsides" rather than their "insides."

Neutral focus arousal. For the physical neutral focus arousal condition, Ss were asked to view a mountain scene for 1 minute, inspect a United States road map for 1 minute, listen to taped instrumental music for 1 minute, and look at a modern painting for 1 minute. In the comparable cognitive condition, Ss were asked to imagine carrying out the above-cited activities for comparable lengths of time.

Arousal Validating Methods

In order to evaluate whether the arousal conditions, described in the preceding section, affected somatic focus, Ss were administered both the Homonym test (Secord, 1953) and the Body Sensation Report technique (Fisher, 1965) immediately after undergoing the arousal manipulations.

The Homonym test is a list of 100 homonyms, 75 of which may have either somatic or nonsomatic meanings (e.g., colon, digit, gag, nail, vessel, etc.). It was hypothesized that if the arousal procedures were effective, the internal and external focus Ss should give more somatic homonym definitions than the neutral focus group. The Body Sensation Report technique involves having an S close his eyes, focus his attention on his body, and report each time a sensation is experienced. The sensations reported by Ss are scored either internal or external depending on the reported body locus. It was hypothesized that if the arousal conditions were adequate for effecting a differential somatic focus, the external focus arousal Ss should report significantly more external body sensations than the internally aroused group. It was also hypothesized that the internal focus arousal Ss should report more internal body sensations than the external arousal group.

Since the Homonym test does not seem to be as obviously concerned with somatic variables as the Body Sensation Report technique, it was always administered first in order not to bias Ss toward bodily responses. Forty-eight Ss (24 male and 24 female) were used for this part of the study.

Rorschach Indexes

After establishing the adequacy of the arousal conditions, a new set of 120 Ss (60 males and 60 females) were used to evaluate the arousal effects on the Rorschach Barrier and Penetration Indexes. The Ss were immediately administered the Rorschach test after they had gone through their particular arousal condition. The protocols were scored for Barrier and Penetration according to the procedures described by Fisher and Cleveland (1958, pp. 58-71). It was hypothesized that if the internal somatic states affect the perceptual organization of unstructured stimuli, then those Ss in the external somatic focus group should have higher Rorschach Barrier scores than do those Ss in the internal and neutral somatic groups. It was also hypothesized that the internal somatic focus Ss would give more Rorschach Penetration responses than Ss in the external and neutral focus conditions.

RESULTS

It was found that the arousal procedures differentially affected the number of homonyms and body sensations reported by the Ss. In addition, Ss in the external somatic focus group gave significantly more Rorschach Barrier responses than the other somatic focus Ss. Significantly higher Rorschach Penetration scores were earned by the internal somatic focus group than by the other Ss.

Scoring Reliability

A random sample of each of the dependent measures was given to a second scorer⁴ who knew nothing about the purpose of the study. Interscorer reliability coefficients (Pearson product-moment correlation) were found to be .95 for the Homonym Test; .87 for both the Internal and External Body Sensation Report Techniques; and .98 and .85 for the Barrier and Penetration scores, respectively.

Arousal Validating Measures

The mean number (and accompanying standard deviations) of somatic homonyms, internal and external body sensation reports, given by the male and female Ss in the physically and cognitively aroused internal, external, and neutral focus groups can be found in the first three columns of Table 1. For the Homonym Test (Column 1 of Table 1), Ss in both the internal and external somatic focus groups produced a higher mean number of somatic homonyms than did Ss in the neutral somatic focus condition (19.25 and 23.69 vs. 15.50). The Ss in the various cognitive arousal conditions gave more somatic homonym definitions ($M = 20.67$) than did those in the physical arousal conditions ($M = 18.29$). The male and female Ss produced similar mean amounts of somatic homonyms (19.21 and 19.75).

In order to evaluate whether any of the mean somatic homonym score differences found among S populations were statistically significant, a $3 \times 2 \times 2$ factorial analysis of

⁴The authors wish to express their gratitude and appreciation to Wynand D. Pienaar for serving as a second protocol scorer.

TABLE 1

MEAN NUMBER OF SOMATIC HOMONYMS, INTERNAL AND EXTERNAL BODY SENSATIONS, AND RORSCHACH BARRIER AND PENETRATION RESPONSES MADE BY THE MALE AND FEMALE SUBJECTS IN THE PHYSICALLY AND COGNITIVELY AROUSED INTERNAL, EXTERNAL, AND NEUTRAL SOMATIC FOCUS GROUPS

Groups	Arousal validating measures (<i>N</i> = 48)						Body boundary indexes (<i>N</i> = 120)			
	Somatic homonyms		Internal body sensation reports		External body sensation reports		Barrier responses		Penetration responses	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Treatment groups										
Male internal physical	16.25	3.6	9.25	1.0	10.75	1.4	3.40	3.0	6.40	2.8
Female internal physical	19.25	7.3	9.00	2.6	8.50	3.4	3.10	1.0	3.60	2.0
Male internal cognitive	21.25	3.4	6.26	2.6	13.75	2.8	3.20	1.4	4.10	2.2
Female internal cognitive	20.25	3.7	9.50	2.2	10.50	2.2	3.30	2.0	6.10	2.4
Male external physical	26.25	12.8	3.75	1.4	16.00	2.2	7.00	1.7	2.40	1.7
Female external physical	17.00	4.2	4.50	1.3	15.50	1.3	6.00	1.4	2.30	1.4
Male external cognitive	24.50	6.1	4.25	1.0	15.75	1.0	6.50	2.4	2.90	2.4
Female external cognitive	27.00	6.7	5.00	1.7	15.00	2.0	5.60	2.2	1.90	1.7
Male neutral physical	14.50	9.1	4.50	3.0	15.50	2.4	2.10	1.4	2.50	1.7
Female neutral physical	16.50	4.3	7.75	2.6	12.25	2.6	2.20	1.0	2.80	1.7
Male neutral cognitive	12.50	3.8	6.25	1.0	13.75	1.0	3.60	1.7	3.40	1.4
Female neutral cognitive	18.50	5.4	4.06	1.4	16.00	1.7	3.10	1.7	2.00	1.7
Arousal procedures										
Physical	18.29	5.2	6.46	1.8	13.08	2.8	3.97	1.8	3.33	1.7
Cognitive	20.67	4.8	5.88	1.9	14.13	2.2	4.22	2.0	3.40	1.8
Somatic focus										
Internal	19.25	3.4	8.50	2.1	10.88	2.2	3.25	1.3	5.05	2.3
External	23.69	6.7	4.38	1.3	15.56	1.8	6.28	2.8	2.38	1.9
Neutral	15.50	5.7	5.63	1.5	14.38	1.9	2.75	1.4	2.68	1.6
Sex differences										
Male	19.21	5.6	5.71	1.7	14.25	1.7	4.30	1.7	3.62	1.7
Female	19.75	4.3	6.63	2.2	12.96	1.7	3.88	1.6	3.12	1.7

variance was carried out.⁵ The summary of the analysis is presented in Table 2. Only the somatic focus main effect was found to be statistically significant ($F = 6.53$, $df = 2/36$, $p < .05$). Further analyses using t tests indicated that the mean number of somatic homonym definitions given by both the internal and external somatic focus Ss were significantly greater than those of the neutral somatic focus Ss. It would appear that the internal and external somatic focus arousal procedures sensitized Ss in such a way as to produce somatic homonym definitions.

⁵ Since the same statistical design was used throughout the study, there will be no further descriptions of the statistical analysis procedures used for the ensuing measures.

An inspection of the mean number of internal body sensation reports (Column 2, Table 1) made by Ss in this study reveals that the internal somatic focus groups gave more internal body sensations ($M = 8.50$) than the external and neutral somatic focus groups ($M = 4.38$ and 5.63 , respectively). The physical arousal condition Ss tended to report more internal body sensations ($M = 6.46$) than did those in the cognitive arousal conditions ($M = 5.88$). A greater mean number of internal body sensations tended to be reported by the female (6.63) than by the male Ss (5.71).

The analysis of variance conducted to reveal whether any of the differences were statistically significant is presented in Table 3.

The somatic focus main effect was significant ($F = 15.22$, $df = 2/36$, $p < .05$). Further analyses revealed that the internal somatic focus Ss generally reported significantly more internal sensations than did the external and neutral somatic focus groups. The Somatic Focus \times Arousal Procedure \times Sex Difference interaction was also found to be significant ($F = 4.37$, $df = 2/36$). It seemed that although the internal somatic focus Ss reported more internal somatic sensations than did the other Ss, the male cognitively aroused internal somatic Ss gave significantly fewer internal body sensations than did the concomitant female group. In addition, the female physically aroused neutral somatic focus Ss reported more internal body sensations than did the other physically aroused groups (excepting the internal somatic focus Ss). All in all, it would seem that the internal somatic focus manipulations led to significantly more internal body sensation reports than did any of the other arousal procedures.

Column 3 of Table 1 presents the mean external sensations reported by the Ss. The external somatic focus Ss reported more external sensations ($M = 15.56$) than did the internal and neutral somatic focus Ss ($M = 10.88$ and 14.38 , respectively). The cognitively aroused groups tended to produce more external sensations ($M = 14.13$) than did the physically aroused Ss ($M = 13.08$). The males also tended to give more external body

TABLE 3
ANALYSIS OF VARIANCE OF THE INTERNAL BODY SENSATIONS REPORTED BY MALES AND FEMALES IN THE PHYSICALLY AND COGNITIVELY AROUSED INTERNAL, EXTERNAL, AND NEUTRAL SOMATIC FOCUS GROUPS

Source	df	MS	F
Somatic focus (SF)	2	71.58	15.22*
Arousal procedure (AP)	1	4.08	.86
Sex differences (SD)	1	10.08	2.14
SF \times AP	2	3.58	.76
SF \times SD	2	1.08	.22
AP \times SD	1	1.33	.28
SF \times AP \times SD	2	20.58	4.37*
Residual	36	4.70	
Total	47		

* $p < .01$.

sensations than did the females ($M = 14.25$ and 12.96 , respectively).

The analysis of variance of the external body sensation scores is presented in Table 4. The somatic focus main effect was found to be significant ($F = 22.25$, $df = 2/36$, $p < .01$). Further analysis disclosed that the external and neutral somatic focus groups reported significantly more external body sensations than did the internal somatic focus Ss. Although the external somatic focus group gave a higher mean external body sensation score than did the neutral somatic focus group (15.56 vs. 14.38), there was no statistically significant difference between the group

TABLE 2

ANALYSIS OF VARIANCE OF THE SOMATIC HOMONYM DEFINITIONS MADE BY THE MALES AND FEMALES IN THE PHYSICALLY AND COGNITIVELY AROUSED INTERNAL, EXTERNAL, AND NEUTRAL SOMATIC FOCUS GROUPS

Source	df	MS	F
Somatic focus (SF)	2	268.77	6.53*
Arousal procedure (AP)	1	67.68	1.64
Sex differences (SD)	1	3.52	.08
SF \times AP	2	18.18	.44
SF \times SD	2	55.02	1.34
AP \times SD	1	46.02	1.12
SF \times AP \times SD	2	62.02	1.51
Residual	36	41.18	
Total	47		

* $p < .05$.

TABLE 4

ANALYSIS OF VARIANCE OF THE EXTERNAL BODY SENSATIONS REPORTED BY THE MALES AND FEMALES IN THE PHYSICALLY AND COGNITIVELY AROUSED INTERNAL, EXTERNAL, AND NEUTRAL SOMATIC FOCUS GROUPS

Source	df	MS	F
Somatic focus (SF)	2	95.02	22.25**
Arousal procedure (AP)	1	13.02	3.05
Sex difference (SD)	1	20.02	4.69*
SF \times AP	2	8.27	1.93
SF \times SD	2	6.39	1.50
AP \times SD	1	6.02	1.41
SF \times AP \times SD	2	12.64	2.96
Residual	36	4.27	
Total	47		

* $p < .05$.

** $p < .01$.

TABLE 5

ANALYSIS OF VARIANCE OF THE RORSCHACH BARRIER SCORES MADE BY THE MALES AND FEMALES IN THE PHYSICALLY AND COGNITIVELY AROUSED INTERNAL, EXTERNAL, AND NEUTRAL SOMATIC FOCUS GROUPS

Source	df	MS	F
Somatic focus (SF)	2	145.50	41.45*
Arousal procedure (AP)	1	1.89	.53
Sex differences (SD)	1	5.20	1.48
SF × AP	2	7.27	2.07
SF × SD	2	2.15	.61
AP × SD	1	.10	.02
SF × AP × SD	2	.65	.19
Residual	108	3.51	
Total	119		

* $p < .01$.

means. The sex difference main effect was also found to be significant ($F = 4.69$, $df = 2/36$, $p < .01$). The male Ss tended to report more external body sensations than did the females. This result is the reverse of the sex difference trend found for the internal body sensation report data. It would seem that the external somatic focus manipulations induced more external body sensation reports than did the other somatic focus procedures.

Rorschach Indexes

The mean Rorschach Barrier and Penetration scores (and accompanying standard deviations) made by the physically and cognitively aroused male and female internal, external, and neutral somatic focus Ss may be found in Columns 4 and 5 of Table 1.

An inspection of Column 4 of Table 1 reveals that the external somatic focus Ss produced higher mean barrier responses than did the other groups of Ss (6.28 vs. 3.25 and 2.75). The cognitively aroused groups tended to give more Barrier responses ($M = 4.22$) than did the physically aroused groups ($M = 3.97$). Little difference in terms of mean production of Barrier responses was found between the sexes (4.30 and 3.88 for the males and females, respectively).

The analysis of variance carried out to establish whether any of the mean differences were statistically significant is presented in Table 5. Only the somatic focus main effect was significant ($F = 41.45$, $df = 2/108$,

TABLE 6

ANALYSIS OF VARIANCE OF THE RORSCHACH PENETRATION SCORES MADE BY THE MALES AND FEMALES IN THE PHYSICALLY AND COGNITIVELY AROUSED INTERNAL, EXTERNAL, AND NEUTRAL SOMATIC FOCUS GROUPS

Source	df	MS	F
Somatic focus (SF)	2	85.90	20.40*
Arousal procedure (AP)	1	.13	.03
Sex differences (SD)	1	7.50	1.78
SF × AP	2	.01	—
SF × SD	2	.07	.01
AP × SD	1	4.03	.96
SF × AP × SD	2	31.40	7.46*
Residual	108	4.21	
Total	119		

* $p < .01$.

$p < .01$). Further analyses indicated that the external somatic focus group gave significantly more Barrier responses on the Rorschach than did the internal and neutral somatic focus Ss. It would seem that external somatic focus manipulations induce Barrier responses to appear on the Rorschach test.

In Column 5 of Table 1 can be found the mean Rorschach Penetration scores made by the Ss in this study. The internal somatic focus groups gave more Penetration responses ($M = 5.05$) than did the internal and neutral somatic focus groups ($M = 2.38$ and 2.68 , respectively). The mean Penetration score differences between the two arousal procedure conditions and between the sexes were rather small (3.33 vs. 3.40, and 3.62 vs. 3.12).

The analysis of variance of these mean differences can be found in Table 6. The somatic focus main effect was found to be significant ($F = 20.40$, $df = 2/108$, $p < .01$). Further analysis indicated that the internal somatic focus Ss gave significantly more Penetration responses than did the external and neutral somatic focus Ss. The Somatic Focus × Arousal Procedure × Sex Difference interaction effect was significant ($F = 7.46$, $df = 2/108$, $p < .01$). Although all the internal somatic focus Ss gave more Penetration scores than did all the other Ss, the male physically aroused internal somatic focus Ss gave more Penetration responses than the cognitively aroused internal somatic focus males. The reverse was true for the female internal somatic

focus Ss.⁶ It appears that the internal somatic conditions induce Penetration responses to appear on the Rorschach test.

DISCUSSION

The results of this study support the proposition that body image boundary status affects the perceptual organization of unstructured stimuli. This proposition has been the prime underlying theoretical assumption behind Fisher and Cleveland's research program relating Rorschach Barrier and Penetration scores to various aspects of personality. They have, however, presented no direct evidence to support this assumption. Without such support, the criticism that the Barrier and Penetration scores are subsidiary to and therefore reflect just another facet of general cognitive-personality style is one that must be taken quite seriously.

However, the present study corroborates Fisher and Cleveland's underlying assumption and mitigates against the Mednick and Wylie skepticism about the relationship between body image processes and perception. If the exterior of the body is made the center of focus, it is reflected in Rorschach response productivity having to do with external somatic locus. When the interiors of the body are experimentally stimulated, there is a concomitant rise in interior physical locus responses on the Rorschach. It is not even necessary to physically stimulate the two body image sectors. These effects are so powerful that even instructions to think about the interior and exterior parts of the body can mediate differential perception on the Rorschach test. There were no cognitive-physical arousal procedure differences found in this study.

The confidence in the confirmation of Fisher and Cleveland's underlying assumption was strengthened by the method used in this study. Attempts to validate the effectiveness

of the independent variable are rare in the research literature. In this study the arousal procedures seemed to successfully induce Ss to focus on the exteriors and interiors of their body. The results can therefore be attributed to the Rorschach's sensitivity to perceptual organization induced by differential somatic focus. The use of an experimental manipulation procedure (rather than a correlational design) helped to demonstrate the existence of a direct relationship between somatic focus and Rorschach responses rather than one in which the two variables co-vary as a function of some other process (such as general cognitive style).

It is difficult to deny a major role to the body image as initiator in the organization of perception and cognition. That bodily states influence perception has already been demonstrated in the work of Clark (1952) and Levine, Chein, and Murphy (1942) on the effects of sex and hunger drives on perception. Somatic influence has also been demonstrated in perception in the sensory-tonic research conducted by Werner and Wapner (1952) and Witkin (1949). It makes sense that the host of physical and cognitive stimulations leading to self-world differentiation affects self-perception and perception of the environment. Weak and penetrated body images projected onto the world can lead to the perception of things as fuzzy, fluid, and open to invasion and insult. Strong body image boundaries should logically lead to perceptions of the world in which contours are articulated and resistant to stress. It is still not known what mediational processes convert body image status into Rorschach Inkblot perception. This teasing out of the processes involved should be the major task of those psychologists interested in the psychosomatic basis of perception.

REFERENCES

- CLARK, R. A. The projective measurement of experimentally induced sexual motivation. *Journal of Experimental Psychology*, 1952, 44, 391-399.
- FISHER, S. A further appraisal of the body boundary concept. *Journal of Consulting Psychology*, 1963, 27, 62-74.
- FISHER, S. Body sensation and perception of projective stimuli. *Journal of Consulting Psychology*, 1965, 29, 135-138.

⁶ These differential sex difference findings, as well as those reported elsewhere (Fisher & Renik, 1966; Renik & Fisher, 1968), may be the result of either learned cultural sex-role patterns of body focus orientation, or endocrinological sex variations, or possibly be due to the effects of the sex of the examiner. Further systematic research is needed to shed light on the sources of sex difference results.

- FISHER, S., & CLEVELAND, S. E. *Body image and personality*. Princeton, N.J.: Van Nostrand, 1958.
- FISHER, S., & CLEVELAND, S. E. Personality, body perception and body image boundary. In S. Wapner & H. Werner (Eds.), *The body percept*. New York: Random House, 1965.
- FISHER, S., & FISHER, R. L. Body image boundaries and patterns of body perception. *Journal of Abnormal and Social Psychology*, 1964, 68, 255-262.
- FISHER, S., & RENIK, O. D. Induction of body image changes. *Journal of Projective Techniques and Personality Assessment*, 1966, 30, 429-434.
- FREUD, S. *The ego and the id*. London: Hogarth, 1927.
- LEVINE, R. I., CHEIN, L., & MURPHY, G. The relation of intensity of a need to the amount of perceptual distortion. *Journal of Psychology*, 1942, 13, 283-293.
- MEDNICK, S. The body's barriers go Rorschach. *Contemporary Psychology*, 1959, 4, 276-277.
- RENIK, O. D., & FISHER, S. Induction of body image boundary changes in male subjects. *Journal of Projective Techniques and Personality Assessment*, 1968, 32, 45-48.
- SECORD, P. F. Objectification of word association procedures by the use of homonyms: A measure of body cathexis. *Journal of Personality*, 1953, 21, 479-495.
- WAPNER, S., & WERNER, H. An experimental approach to body perception from the organismic-developmental point of view. In S. Wapner & H. Werner (Eds.), *The body percept*. New York: Random House, 1965.
- WERNER, H., & WAPNER, S. Towards a general theory of perception. *Psychological Review*, 1952, 59, 324-338.
- WITKIN, H. A. Perception of body position and the position of the visual field. *Psychological Monographs*, 1949, 63(7, Whole No. 302).
- WITKIN, H. A. Development of the body concept and psychological differentiation. In S. Wapner & H. Werner (Eds.), *The body percept*. New York: Random House, 1965.
- WYLLIE, R. C. *The self concept*. Lincoln: University of Nebraska Press, 1961.

(Received August 15, 1968)

SUICIDE AND NATURAL DEATH IN A STATE HOSPITAL POPULATION:

A COMPARISON OF ADMISSION COMPLAINTS, MMPI PROFILES, AND SOCIAL COMPETENCE FACTORS

MILTON R. RAVENSBORG¹ AND ADELINE FOSS

Fergus Falls State Hospital, Minnesota

Completed suicides in a state hospital were contrasted with natural deaths and inpatients on admission complaints, MMPI profiles, and social competence factors. No significant differences were found among the groups on patterns of admission complaints. For all three groups, depression was found to be the most frequent admission complaint. Neither MMPI profiles nor a recently developed suicide threat scale served to discriminate suicides from the comparison groups. Highly significant differences were found on social competence variables, with suicides tending to be younger, more intelligent, and better educated than comparison groups.

A review of the literature suggests that there are three major approaches to the study of suicide: the clinical, the psychometric, and the sociological (Farberow, Shneidman, & Neuringer, 1966). Each of these approaches has yielded valuable insights to the concomitants of suicide. However, as Kahne (1966) pointed out, no approach has provided a royal road to understanding; each is a "style of explanation."

The clinical approach, based on the symptoms and personality traits of identified suicidal cases, has resulted in such conclusions as "depression is the major suicidal illness [Shabshin, 1965]." However, Eisenthal, Farberow, and Shneidman (1966), after a detailed follow-up of suicide status patients of a Veterans Administration Hospital, reported, "Our data did not show that the severity of depression was related to suicide potentiality. Patients with psychotic depressions were less suicidal than most other psychotics and neurotics [p. 988]."

The psychometric approach, based on test items and other indicators produced in the course of psychological testing, has resulted in scales which are alleged to discriminate between suicidal persons and others (e.g., Farberow & Devries, 1967; Holzberg, Cohn, & Wilk, 1951). However, such scales and

indexes fail sometimes to operate when applied to new populations, or in settings which differ from those in which such indexes were originally validated.

The sociological approach, based on social and economic factors presented by suicidal persons and others (e.g., Farberow et al., 1966) has so far had little impact on the mainstream of psychiatric thought about suicide. Kahne (1966) urged suicide researchers not to concentrate solely on predisposing factors within the patient, and chided physicians for their disregard of sociological and anthropological contributions.

Because clinical, psychometric, and sociological experts tend to prefer their own styles of explanation, one rarely encounters studies which view suicidal phenomena from all three standpoints. As a result, it is impossible to ascertain, within any given population, whether the clinical symptom picture, the psychometric profile, or the background of socioeconomic factors is the most useful in discriminating suicidal patients from control groups.

One of the few investigations of suicide which used the three approaches (clinical, psychometric, and sociological) at once was that of Simon (1950), who analyzed a sample of 50 veterans who attempted suicide. However, Simon failed to employ a relevant control group (e.g., other hospitalized veterans). Thus his results, while interesting,

¹ Requests for reprints should be sent to Milton R. Ravensborg, Fergus Falls State Hospital, Fergus Falls, Minnesota 56537.

did not lead to indexes which would help in identifying potential suicides.

The present study was undertaken to learn whether the patterns of admission complaints, psychometric (MMPI) profiles, or social competence factors were most efficient in discriminating suicides from other psychiatric patients.

METHOD

From Fergus Falls State Hospital records, the complete roster of on-books psychiatric patients who had died in the period January 1, 1954, through June 30, 1967, was obtained ($N=275$). On the basis of death certificate information, this roster was divided into two groups, patients who died by suicide ($N=27$) and patients who died of natural causes. The list of patients who died natural deaths was then scrutinized, and the two largest subgroups, patients who died from circulatory and respiratory diseases ($N=173$), were culled and combined as the natural death population. The third group in the investigation, regular psychiatric patients, was the current in-hospital population of a geographical psychiatric unit ($N=109$).

Psychometric files were obtained for all Ss in the three populations, for whom such files were available. MMPIs and Shipley-Hartford scales were available for 23 (85%) of the suicides (15 males, 8 females); 23 (13%) of the natural deaths (15 males, 8 females); and 60 (55%) of the inpatients (28 males, 32 females). It should be noted that these percentages of "testable" Ss are significantly different ($\chi^2=116.13$, $p<.001$). This finding will be treated in the Discussion section.

Because of the small N s in the two death groups, males and females were combined and, to enable use of the profile analysis of variance (Block, Levine, & McNemar, 1951), a random sample of 23 inpatients was drawn.

MMPI profiles were tabulated separately for the three groups, combining males and females. (Subsequent comparisons are restricted to "testable" Ss: those for whom all types of data considered in this report were available.) Group profiles were then analyzed by analysis of variance. In addition, a recent Suicide Threat scale (Farberow & Devries, 1967) was scored on the three groups.

Admission diagnosis, based on the American Psychiatric Association's two-digit code (American Psychiatric Association, 1952) was obtained for each S, and the frequency of each diagnostic type was tabulated for the three groups. Frequency arrays were then compared by the chi-square test of independence.

Median time-in-hospital was calculated for each of the three groups and the groups were compared with each other by the extension of the median test (Siegel, 1956).

A schedule of admission complaints was derived by perusing admission records of the 109 inpatients and listing every admission complaint found in any

patient's file. Individual patient schedules were then summarized for this group to obtain the frequency of each separate complaint. All complaints which occurred five or more times in the population of 109 were included in the final schedule, which contains 34 admission complaints.

Using the admission complaint schedule, admission notes in the charts of the 23 "testable" suicidal patients and the sample of 23 "testable" natural deaths were analyzed. The frequency of each of the 34 complaints in the two deceased populations was enumerated. Likewise, the frequency of each complaint was determined for the random sample of 23 inpatients. Rank orders of the admission complaints were determined (Table 1), and Kendall's Coefficient of Concordance (Siegel, 1956) was computed to determine the relatedness of the rank orders.

Social competence factors were coded for the three populations on the basis of the Zigler and Phillips (1961) scale (Table 2). The social competence scale has six factors (age, intelligence, education, occupation, employment history, and marital status). Each factor is scored 0, 1, or 2, with 2 representing higher social competence. In the present study, education was regrouped, with less than high school diploma = 0; high school diploma = 1; education beyond high school = 2. The portions of each population which fell into the three levels of each social competence factor were expressed as percentages.

RESULTS

Diagnosis

The various diagnostic types (based on the American Psychiatric Association's two-digit code) were represented in about equal proportions in the three groups ($\chi^2=2.83$, $p>.50$). About half of the patients in each group had been diagnosed as suffering from some form of schizophrenic reaction (Table 3).

The median time-in-hospital for inpatients was 22 months (range: 4-366); for suicides, 17.5 months (range: 6-231); for natural deaths, 23.0 months (range: 1-356). The overall median for the three groups, combined, was 23.0 months. The extension of the median test (Siegel, 1956) resulted in a chi-square of .34 (ns), implying that there were no significant differences among median lengths-of-stay for the three groups.

Admission Complaints

There was no significant difference in rank order of admission complaints among the three populations of suicides, natural deaths, and psychiatric inpatients (Table 1). Kendall's Coefficient of Concordance (Siegel,

1956) yielded a W of .63, equivalent to chi-square 62.37 ($p < .001$). Of particular interest is the fact that "depressed" was the top-ranking complaint for all three groups.

MMPI Profiles

The analysis of variance of MMPI profiles yielded an $F = 2.09$. Following Greenhouse and Geisser (1959), degrees of freedom was adjusted from 28 and 924 (Interaction \times Residual) to 2 and 66 (Groups \times Individuals) because of the lack of independence of the scale variables. The obtained F was not significant ($p > .05$). Thus, on the basis of present evidence, it does not appear feasible to distinguish suicides from other patients on the basis of MMPI profiles.

Application of the Suicide Threat scale to MMPI answer sheets of the three groups of Ss yielded means of 21.65 for suicides and 22.93 for other patients. Analysis of variance included a breakdown by sexes. F ratios for main effects did not achieve significance, implying that suicides, as a whole, did not differ from other patients on this scale; neither did males, as a group, differ from females. The Sex \times Patient Type interaction, however, was significant ($p < .025$), due to the fact that female inpatients achieved the highest mean of all. These results suggest that the Suicide Threat scale was of little value in differentiating the state hospital population.

TABLE 1
RANKINGS OF ADMISSION COMPLAINTS

Complaint	Inpatient	Suicide	Natural death
Depressed	1	1	1
Suspicious	2	9.5	2
Hallucinations	3	9.5	9
Delusional	4.5	3.5	4.5
Nervous, overactive	4.5	18	19
Aggressive and delinquent behavior	6	6.5	19
Abusive and angry	7	9.5	19
Preoccupations	8	26.5	25.5
Seclusive	9	18	13.5
Threatened injury to others	10	13.5	6.5
Insomnia	11	2	25.5
Loss of interest	12	22.5	13.5
Threatened or attempted suicide	13	6.5	25.5
Withdrawn	14	13.5	25.5
Disoriented	15.5	31.5	25.5
Somatic complaints	15.5	5	9
Confused	17	9.5	4.5
Overly dependent, unable to care for self	19	22.5	32
Silly	19	31.5	19
Violent emotional outbursts	19	26.5	6.5
Forgetful	19	22.5	3
Feelings of persecution	22	26.5	9
Fearful	23.5	13.5	13.5
Excessive drinking	23.5	18	13.5
Bizarre behavior	25.5	31.5	13.5
Feelings of rejection	25.5	31.5	32
Inferiority feelings	27.5	18	25.5
Guilt feelings	27.5	13.5	25.5
Abnormal sex behavior	29	31.5	32
Inability to work	30.5	18	13.5
Irritable	30.5	22.5	25.5
Moody	32	26.5	32
Phobias	33	31.5	32
Anxious	34	3.5	19

Note.—Kendall's Coefficient of Concordance of rank orders significant beyond the .001 level ($\chi^2 = 62.37$).

Social Competence

Considerable variance among the three populations was present in the tabulation of social competence factors (Table 2). All chi-square comparisons were significant beyond the .001 level.

In regard to age, it was evident that inpatients tended to be somewhat older than suicides—18% of the suicides were in the age group of under 25, whereas only 5% of the inpatients and none of the natural deaths fell into this age bracket.

As to intelligence, 27% of the suicides were above IQ 115, whereas only 8% of the inpatients and none of the natural deaths had tested at this intellectual level.

Fourteen percent of the suicides had re-

ceived education beyond the high school diploma, while only 3% of the inpatients and 9% of the natural deaths had studied beyond high school. On the other hand, 94% of the inpatients and 82% of the natural deaths had failed to complete high school.

In regard to occupation, the vast majority of all three populations were coded at the unskilled or semiskilled levels. However, 9% of the suicides had held jobs at the professional and managerial level, whereas only 1% of the inpatients, and none of the natural deaths, had progressed to such top-level jobs.

With respect to employment history, more of the natural deaths had been usually unemployed and fewer had been employed on an irregular basis. There was little evidence, however, that suicides could be differentiated from

TABLE 2
SOCIAL COMPETENCE-CODING SYSTEM AND PERCENTAGE OF INPATIENTS,
SUICIDES, AND NATURAL DEATHS IN EACH CATEGORY

Variables	Code	Populations ^a		
		IP	S	NS
Age	0	5	18	0
24 and below	1	22	46	36
25-44	2	73	36	64
45 years and above				
Intelligence	0	25	5	23
84 or less	1	67	68	77
85-115	2	8	27	0
116 and above				
Education				
None or some grades including ungraded or special classes, and some high school	0	94	50	82
High school diploma, no more	1	3	36	9
Education beyond high school: some college, or vocational training	2	3	14	9
Occupation	0	77	55	86
Unskilled or semiskilled	1	22	36	14
Skilled, service, clerical, sales	2	1	9	0
Professional and managerial				
Employment history	0	16	14	32
Usually unemployed				
Seasonal, fluctuating, frequent shifts, or part-time employment	1	56	50	32
Regularly employed	2	28	36	36
Marital status				
Single	0	75	40	23
Separated, divorced, remarried, or widowed	1	16	14	32
Single continuous marriage	2	9	46	45

Note.—All N's = 23. IP = Inpatients, S = Suicides, ND = Natural deaths.
^a Chi-squares of the percentages of the three populations falling into each of the three levels of the six Social Competence factors ranged from 17.47 to 66.10, all significant beyond .001.

TABLE 3

FREQUENCY OF DIAGNOSTIC TYPES IN INPATIENTS, SUICIDES, AND NATURAL DEATHS

Diagnostic type	Group		
	IP	S	ND
Chronic brain syndrome	1	2	5
Psychotic disorders			
Affective reaction	5 ^a	2 ^b	1 ^c
Schizophrenic reaction	11	12	10
Psychoneurotic disorders	1 ^c	3 ^a	4 ^b
Personality disorders	3	4	3
Mental deficiency	2	0	0
Total	23	23	23

Note.—IP = Inpatients, S = Suicides, ND = Natural deaths, $\chi^2 = 2.83$, $p > .50$.

^a Includes three cases of depressive reaction.

^b Includes two cases of depressive reaction.

^c Includes one case of depressive reaction.

inpatients or natural deaths on the basis of employment history.

As to marital status, 75% of the inpatients were single, whereas only 23% of the natural deaths had failed to marry. Forty-six percent of the suicides has had a single continuous marriage, as did 45% of the natural deaths. By contrast, only 9% of the inpatients had been steadily married to one partner.

DISCUSSION

Because of the very low percentage of "testable" natural deaths, and the relatively low percentage of inpatients whose files contained MMPIs and Shipley-Hartford scales, the conclusions of this study should probably not be generalized to all members of the inpatient and natural death populations. But, inasmuch as 85% of the suicides had been tested, the obtained mean MMPI profile, pattern of admission complaints, and social competence factors might be considered representative of the suicide group.

The finding that a significantly larger percentage of suicidal patients had been found intact enough to complete MMPIs in a group-testing situation adds support to the inference, based on social competence comparisons, that suicides tend to be more intelligent and better educated (as well as younger) than inpatients and natural deaths.

On the basis of the present investigation, it is evident that no pattern of admission complaints, let alone the single complaint, "depression," is particularly useful in identifying potential suicides among psychiatric patients.

Likewise, the MMPI profile appears to add nothing to discrimination of self-destructive patients as a class. A comparable conclusion, based on multiple administrations of the MMPI to five suicidal patients, was recently published by Devries and Shneidman (1967). The authors speculated that "the establishment of suicidal MMPI profile syndromes can best be accomplished if persons with homogeneous rather than . . . heterogeneous syndromes are grouped together for comparison purposes [p. 401]."

However, the worth of Kahne's (1966) admonitions to look at sociological factors in suicide research was borne out by the finding of significant differences in social competence between suicides and other state hospital patients. It is not, however, immediately apparent why the suicides in this population should present a different pattern of background factors from those of inpatients and natural deaths.

REFERENCES

- AMERICAN PSYCHIATRIC ASSOCIATION. *Diagnostic and statistical manual*. Washington, D. C.: Author, 1952.
- BLOCK, J., LEVINE, L., & McNEMAR, Q. Testing for the existence of psychometric patterns. *Journal of Abnormal and Social Psychology*, 1951, 46, 356-359.
- DEVRIES, A. G., & SHNEIDMAN, E. S. Multiple MMPI profiles of suicidal persons. *Psychological Reports*, 1967, 21, 401-405.
- EISENTHAL, S., FARBEROW, N. L., & SHNEIDMAN, E. S. Follow-up of neuropsychiatric patients in suicide observation status. *Public Health Reports*, 1966, 81, 977-990.
- FARBEROW, N. L., & DEVRIES, A. G. An item differentiation analysis of MMPIs of suicidal neuropsychiatric hospital patients. *Psychological Reports*, 1967, 20, 201-217.
- FARBEROW, N. L., SHNEIDMAN, E. S., & NEURINGER, C. Case history and hospitalization factors in suicides of neuropsychiatric hospital patients. *Journal of Nervous and Mental Disease*, 1966, 142, 32-44.

- GREENHOUSE, S., & GEISSER, S. On methods in the analysis of profile data. *Psychometrika*, 1959, 24, 95-112.
- HOLZBERG, J. D., COHNE, E. R., & WILK, E. K. Suicide: Psychological study of self destruction. *Journal of Projective Techniques*, 1951, 15, 339-354.
- KAHNE, M. J. Suicide research: A critical review of strategies and potentialities in mental hospitals. *International Journal of Social Psychiatry*, 1966, 12, 120-129.
- SHABSHIN, M. The treatment of depression—the major suicidal illness. In L. Yochelson (Ed.), *Symposium on suicide*. Washington, D. C.: George Washington University School of Medicine, 1965.
- SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
- SIMON, W. Attempted suicide among veterans. *Journal of Nervous and Mental Disease*, 1950, 111, 451-468.
- ZIGLER, E., & PHILLIPS, L. Social competence and outcome of psychiatric disorder. *Journal of Abnormal and Social Psychology*, 1961, 63, 264-271.

(Received August 15, 1968)

FEMALES' AFFECTIVE RESPONSES TO READING EROTIC LITERATURE

DONALD L. MOSHER¹

University of Connecticut

AND

IRENE GREENBERG

Ohio State University

The affective states of sexual arousal, guilt, anxiety, and 11 Nowlis mood factors were studied in 72 college females before and after reading either an erotic or an academically oriented literary passage, with a female *E* either present or absent, and as a function of the personality disposition of sex guilt. Females who read the erotic passage showed a significant increase in sexual arousal. High sex guilt females significantly increased their affective state of guilt when they read the erotic passage in comparison to the females in all of the other cells. Anxiety increased as a function of reading the erotic passage in the presence of *E*. Conceptual distinctions between sexual behavior and sexual arousal, between guilt as personality disposition and as state, and between affective states of fear and guilt were discussed.

Kinsey and his associates (Kinsey, Pomeroy, Martin, & Gebhard, 1953) in their study of sexual behavior in the human female pointed out major sex differences in response to erotic stimuli. In their surveys which included approximately 4,000 males and 5,700 females, they found that the males were more responsive than the females to 29 of 33 categories of stimuli that were potentially erotic. One of the three classes of sexual stimuli to which females were as responsive as males was the reading of erotic literary material. In their sample of 5,699 women, 16% reported definite and/or frequent erotic response to reading literary material, 44% reported some erotic response, and 40% reported never experiencing an erotic response to reading literary material (Kinsey et al., p. 670). The present experiment investigated the affective responses of college women to reading erotic literary material as a function of individual differences in the personality disposition of sex guilt and response to the presence of a female *E*. Reports of increased states of sexual arousal, guilt, and anxiety were the major dependent variables.

The major hypothesis under investigation was that high sex guilt females will report a greater increase in the affective state of guilt after reading erotic literature than will low sex guilt females. Violation of prohibi-

tions concerning sexual behavior are expected to lead to the experience of the state or feeling of guilt in guilt-disposed individuals. Okel and Mosher (1968) demonstrated that males scoring above the median on a measure of guilt over hostility reported experiencing a significantly greater affective state of guilt after they had verbally attacked a stranger who became quite upset than did low-hostility-guilt males. The high-guilt males also decreased significantly in their report of positive affects after guilt over aggression had been induced in the experimental situation. Also, it seemed plausible that high sex guilt females may report less sexual arousal than low sex guilt females. Perhaps, those women in Kinsey's sample who reported no sexual arousal to reading erotic literature experienced guilt-produced conflict which inhibited the development of erotic feelings and fantasies.

This experiment also explored the affective reaction of the women as a function of the presence or absence of a female *E*. The presence of an observer might tend to inhibit the arousal of sexual feelings and fantasies or to increase anxiety or fear. The purpose of this manipulation was to add further empirical evidence to support the feasibility of making a conceptual distinction between fear and guilt. Fear is a situational expectancy for punishment that may be increased by the presence of another person who might potentially censure the individual for sex-related behav-

¹ Requests for reprints should be sent to Donald L. Mosher, Department of Psychology, University of Connecticut, Storrs, Connecticut 06268.

ior. Guilt is a generalized expectancy for self-mediated punishment for violating internalized standards concerning sexual behavior. A person feels guilty when he violates his moral standards regardless of the presence or attitudes of others. Mosher (1965) found that males who were low on sex guilt were more influenced by situational cues concerning censure from an *E* than were high sex guilt males in responding to a perceptual defense task. Galbraith and Mosher (1968) found that both situationally manipulated expectancies for *E* censure and sex guilt predicted the inhibition of sexual associative responses to *double entendre* words in a word association task. In this experiment it was anticipated that females will report more of an increase in fear or anxiety—but not in guilt—when they read erotic literature in the presence of *E* than when alone.

METHOD

Procedure

The Ss were 72 females enrolled in the introductory psychology course at the Ohio State University. The Ss were seen twice. Three groups totaling 85 Ss were administered the female form of the Mosher Forced-Choice Guilt Inventory. The Ss were subsequently divided at the median of 18 to form high- and low-guilt groups and randomly assigned to four experimental treatments in which they participated during a second individual experimental session. The Ss were partitioned into eight cells with nine Ss in each cell on the basis of high or low guilt, reading an erotic or an academic literary passage, and *E* presence or absence. The *E* was also female. All Ss were given several opportunities to withdraw from the experiment, and one S declined to read the erotic passage. Twelve Ss of the original 85 were not included in the final sample because they either missed their individual appointments or improperly filled out their inventory or checklists.

After arriving for her individual experimental session, all Ss completed the Nowlis Mood Adjective Check List (MACL) which also included seven guilt adjectives and seven sex-arousal adjectives as a pretest of their affective state at that moment. Then they were instructed, "Now please read the following paragraphs. When you have finished again fill out the MACL, being particularly attentive to your current feelings." The Ss were then given either an erotic or an academic passage to read. The *E* then indicated whether she would be present or absent by saying either, "I'll wait here until you are finished with both" or "I'll wait outside. When you are finished with both just call me in." After the reading of the passage and the posttest MACL were completed, the Ss were debriefed and thanked.

The erotic passage was Chapter 72 of the novel *Eternal Fire* by Calder Willingham (1964). The passage describes the seduction of a virginal girl by a young man. Sexual petting and foreplay lead to mounting sexual arousal in the girl, and her conflicts about sexual intercourse give way until she actively encourages the young man to have sexual intercourse with her. The literary style is vivid and realistically descriptive. The passage was selected because of its highly erotic nature and because it would seem to provide the college female with a heroine with which to identify.

The academically oriented passage was taken from pages 714-721 of the text *Psychology* by Wickens and Meyer (1963). The passage discusses the associative cortex and its functions. The passage was taken from the text in use in Ss' introductory psychology course, but it was unlikely that the Ss would have progressed that far in their reading of their text book. Both passages contained the same number of words, but, of course, the style of writing as well as the content was very different.

Sex-Guilt. The Sex-Guilt subscale (Mosher, 1968) of the female form of the Mosher Forced-Choice Guilt Inventory was constructed from an item pool of 276 completions to the Mosher Incomplete Sentences Test (Mosher, 1961). An internal consistency item analysis and the elimination of items pulling greater than 75% of the responses in a single direction reduced the item pool to 110 items. From this reduced item pool 39 forced-choice sex guilt items were selected. The items are weighted along a guilt dimension from very guilty (+2), guilty (+1), non-guilty (-1), to very nonguilty (-2). The items have a possible range from +64 to -61 and in this sample the obtained range was from +50 to -49. Sample items are:

- Sex relations before marriage . . .
 - A. should be permitted. (SG - 2)
 - B. are wrong and immoral. (SG + 2)
- If I had sex relations, I would feel . . .
 - A. very dirty. (SG + 2)
 - B. happy and satisfied. (SG - 2)

The Sex-Guilt subscale has a corrected split-half reliability of .95. A multitrait-multimethod matrix analysis and a factor analysis based on the responses of 62 female Ss to three measures of three aspects of guilt, including the forced-choice measure of sex guilt, provided evidence for convergent and discriminant validity. The forced-choice sex guilt subscale (FCSG) is not correlated significantly with either the Marlowe-Crowne Social Desirability scale or the Edwards Social Desirability scale. The FCSG subscale is viewed as measuring the personality disposition of guilt rather than the feeling state of guilt.

Nowlis MACL. The Nowlis MACL, which is composed of a number of adjectives applicable to momentary moods or affective states, acts as a series of prompts for statement of mood on the part of the individual. Everyone learns to label his conscious mood states with thousands of adjectives,

and Nowlis believes these descriptive labels prompt a literal "self-report" which is a trustworthy index of present affective states. The aim of Nowlis' research has been to "induce definite changes in mood with clear-cut operations and to determine the relationship between these operations and verbal reports of mood [Nowlis & Green, 1964, p. 2]." The MACL has proven useful in a number of studies which have evaluated changes in mood as a function of such variables as stress, arousal, and drug administration (Cameron & Wendt, 1964; Lazarus & Alfert, 1964; Lazarus, Speisman, Mordkoff, & Davison, 1962; Nowlis, 1965; Nowlis & Green, 1964; Nowlis & Nowlis, 1956; Robbins, 1962).

The form of the MACL used in the present study was composed of 49 adjectives representing 12 factors isolated in a study of mood change of films (Nowlis & Green, 1964). Nowlis and Green report retest reliability coefficients for individual words ranging from .52 to .80. They conclude that "social desirability status of a word has very little, if any, effect on how it is checked when the subject is asked to report how he feels at the moment he reads each word [p. 2]." In addition to the MACL adjectives, seven guilt and seven sexual arousal adjectives were employed. The adjectives used to measure the state of sexual arousal were titillated, sensuous, aroused, tantalized, passionate, excited, and hot. The seven guilt adjectives, which had been used originally by Haefner (1956), were contrite, repentant, ashamed, blameworthy, guilty, conscience stricken, and remorseful. Okel and Mosher (1968) provided construct validation evidence for the usefulness of these adjectives as a measure of the momentary state of guilt in their experiment on changes in affective state as function of induced guilt over aggression. Each adjective could be responded to as definitely applies (4), slightly applies (3), undecided (2), and definitely does not apply (1).

RESULTS

Sexual arousal, guilt, and the 12 affective factors of Nowlis were analyzed by a four-way analysis of variance with a repeated measure on one of the factors. Sex guilt (A), type of passage (B), and *E*'s presence or absence (C) were the between-group factors and the pretest and posttest adjective checklist (D) was the repeated factor.

Table 1 presents the results of the four-way analysis of variance of sexual arousal.² There was a significant interaction ($B \times D$) between reading the sexual passage and an increase in sexual arousal. The Ss who read the erotic passage reported an increase in sexual arousal which was significantly different, using the

² The authors wish to acknowledge the assistance of the Computer Center of the Ohio State University.

TABLE 1
ANALYSES OF VARIANCE OF THE AFFECTIVE STATES OF
SEXUAL AROUSAL, GUILT, AND ANXIETY

Source	df	Sexual arousal	Guilt	Anxiety
		F	F	F
Sex guilt (A)	1		7.93**	
Type of passage (B)	1			4.85*
<i>E</i> 's presence (C)	1			
A \times B	1			
A \times C	1			
B \times C	1			
A \times B \times C	1			
Ss within groups	64			
Adjective check list (D)	1	8.64**		10.01**
A \times D	1		5.03*	
B \times D	1	8.95**	10.61**	
C \times D	1		7.56**	
A \times B \times D	1		7.56**	
A \times C \times D	1			5.14*
B \times C \times D	1			
A \times B \times C \times D	1			
D \times Ss within groups	64			

* $p < .05$.

** $p < .01$.

Newman-Keuls test,³ from their pretest level and both the pretest and posttest levels of the Ss who read the academic passage. The significant main effect of the repeated measure (D) merely reflects this overall increase in sexual arousal as a function of the significant B \times D interaction.

Table 1 also reports the results of a four-way analysis of variance of the guilt adjectives. The major hypothesis was supported since the predicted interaction ($A \times B \times D$) between sex guilt, type of passage, and change in the state of guilt was significant. The mean of the second report of feelings of guilt for the high sex guilt Ss who read the erotic passage was significantly higher than the means of the other seven cells including their own pretest scores. From Table 1, it can be seen that there were three significant two-factor interactions in addition to the significant triple interaction. The A \times D and B \times

³ All significance tests between cell means after significant *F* ratios were obtained used the Newman-Keuls method. All significance levels reported were two-tailed.

D interactions reflect the effects of the $A \times B \times D$ interaction. In the $A \times D$ interaction high sex guilt Ss reported significantly more feelings of guilt on the second MACL than did Ss in the other three cells including their own pretest scores. Also, high sex guilt Ss checked more guilt adjectives on the first administration of the MACL than did the low sex guilt Ss at either time. In the $B \times D$ interaction Ss who read the erotic passage reported a significantly higher mean on the guilt factor on the second administration of the MACL than did Ss in the other three cells. The significant $C \times D$ interaction is a function of a significant increase in the second MACL over the pretest MACL for Ss who read either passage in the presence of *E*. Note that this interaction was not specific to either the erotic passage or to sex guilt. The significant main effect on the repeated measure was a function of an overall decrease in reported guilty state on the second administration of the MACL. While the high sex guilt Ss who read the erotic passage increased in guilt feelings, the other groups all declined in reported guilty state so that there was a significant overall decrease in the mean score on guilty state from the first to the second administration of the MACL.

Table 1 also contains the analysis of the state of anxiety. The significant $B \times C \times D$ interaction for anxiety and the absence of a $B \times C \times D$ interaction for guilt support the feasibility of the conceptual distinction between fear and guilt. The Ss who read the erotic passage with *E* present showed an increase in the mean state of anxiety which was significantly greater than that of Ss in the other seven cells which includes their own first MACL anxiety scores. However, there was a significant increase in the state of anxiety in Ss who read the erotic passage when *E* was not present as well. Note that it was the case, however, that Ss reading the erotic passage with *E* present reported more of an increase in anxiety than Ss reading the erotic passage with *E* absent. The significant *B* main effect indicated that Ss reading the erotic passage were more anxious than Ss reading the academic passage. The *D* main effect indicated greater anxiety in the second MACL.

Both of these main effects are secondary in interpretive significance to the $B \times C \times D$ interaction.

Other Affective States

The results on the other 11 Nowlis mood factors will be briefly summarized. Surgency ($p < .05$), Elation ($p < .01$), and Social Affection ($p < .05$) all decreased significantly on the second administration as simple main effects. For the Egotism factor, there was a significant $B \times C \times D$ interaction ($p < .05$) which is uninterpretable since high sex guilt Ss who were to read the neutral passage, but who did not know it, had a higher mean Egotism score than did three other cells including high sex guilt Ss who were unaware that they were about to read the erotic passage. Egotism ($p < .01$) decreased significantly on the second administration of the MACL. Then, these four positive affective factors declined as a function of participating in the experiment. On the Aggression factor, there was a significant $B \times D$ interaction ($p < .05$) such that Ss who read the erotic passage with *E* present were most angry while Ss who read the academic passage with *E* present were least angry. On the Sadness factor, there was a significant $B \times D$ interaction ($p < .01$) such that Ss who read the erotic passage reported more sadness on the second MACL than Ss in the other three cells. There was also a significant $C \times D$ interaction ($p < .05$) such that Ss reported more sadness on the second MACL when *E* was present. A simple *D* main effect ($p < .01$) reflected the increase in sadness on the second MACL. On the Concentration factor, there was a significant $B \times D$ interaction ($p < .05$) such that Ss who read the academic passage increased in concentration while Ss reading the erotic passage decreased in concentration on the second MACL. On the Nonchalance factor, there was a significant $C \times D$ interaction ($p < .05$) such that Ss who read either passage with *E* present reported feeling less nonchalant on the second MACL; the significant *D* main effect ($p < .01$) reflects this overall decrease in nonchalance. On the Fatigue factor, there was a significant $A \times B \times C$ interaction ($p < .05$) such that high sex guilt Ss are more fatigued with *E*

present while reading the erotic passage and more fatigued with *E* absent while reading the academic passage while the exact reverse holds for the low sex guilt Ss. Fatigue decreased on the second MACL as a main effect ($p < .01$). On the factor of Vigor, there is an uninterpretable $A \times B \times D$ interaction ($p < .05$) in which high sex guilt Ss report a greater feeling of vigor on the first MACL preparatory to reading the academic passage than is the case in three other cells including high sex guilt Ss who are yet to be told they will be reading an erotic passage. There were no significant differences on the Skepticism factor.

DISCUSSION

The data on sexual arousal indicated that females who read the erotic passage reported greater sexual arousal than did the females reading the academic passage. The anticipated inhibitory effect on sexual arousal of the personality disposition of sex guilt was not substantiated; neither did the presence of the female *E* inhibit sexual arousal for those females who read the erotic passage. It may be that these variables require a more complex conceptual structure before relationships, if they exist, can be demonstrated. It may be that sex guilt inhibits sexual arousal when sexual cues are minimal, but there is a threshold of erotic cues such that sexual arousal will occur regardless of the female's disposition toward sex guilt. If this is the case, however, a simple dimension of erotic realism, according to the Kinsey (1953) data, will not suffice for females. Love and sex are more inextricably intertwined in females than in males. In the erotic passage that was used, when it is taken out of the novel's context, the sexual descriptions were accompanied by indications of apparent tenderness and concern. Further research is required to discover the precise roles of erotic realism and the male's affection, tenderness, and commitment on the sexual arousal of females reading erotic literature. Another possible explanation of the failure to find either guilt over sex or *E*'s presence inhibiting sexual arousal might be that these variables inhibit the expression of sexual behavior but not sexual fantasies or sexual arousal. If 40% of the

females in the Kinsey sample report never experiencing erotic arousal while reading literary passages, is it because erotic literature fails to arouse them or is it because they avoid reading erotic literature?

High sex guilt females who read the erotic passage reported an increase in their feeling of guilt which was significantly greater than their pretest state of guilt and that of the remaining females at any time. While high sex guilt females were more likely to report feeling more guilty when they entered the experimental situation, they did significantly increase their feelings of guilt when they read the erotic passage. The affective state described as "feeling guilty" presumably follows transgressions of a moral code. Is the moral violation in this instance that of reading erotic literature as part of an experimental task? Despite repeated reminders that these Ss could decline to participate in this study at any time, only one chose not to continue after reading a few lines of the erotic passage. Or, is the state of guilt engendered in females who are predisposed to guilt over sex when they experience sexual arousal by actively participating in a sexual fantasy while reading the erotic passage? Further research is required before a definitive answer can be given. It does seem helpful to make conceptual distinctions between sexual behavior, thoughts and fantasies, and affects, and between guilt as an inhibitor of behavior and as an affective state which follows a sexual transgression which is either behavioral or cognitive-affective.

Females reported feeling significantly more "fearful," "clutched-up," and "jittery" after reading the erotic passage when *E* was present than did the other females in different treatments or than they had before they read the passage. This significant triple interaction and the failure to find any significant triple or four-way interactions between high sex guilt, reading the erotic passage, with *E* present, and increases in negative affective states support the feasibility of making distinctions like those reviewed in the introduction between affective states such as fear and guilt. However, it should be noted that the females reported feeling an increase in guilt when *E* was pres-

ent regardless of the type of passage read, and that the females were significantly more anxious when they read the sexual passage with *E* out of the room than they were before or than were the females who read the academic passage. These are not crucial weaknesses to the argument, however, since the increase in guilt is not specific to reading the sexual passage and the females who read the erotic passage with *E* present were significantly more anxious than those who read the erotic passage with *E* absent.

The data from the other Nowlis mood factors revealed that females have less positive affects after participating in an experiment of this sort. Reading erotic literature rather than academic literature, in addition to increasing sexual arousal and guilt, increases feeling of sadness while decreasing concentration. While they are reading either passage if *E* is present, in addition to feeling more anxious and guilty, the females feel more sadness and less nonchalance than if *E* waits outside.

REFERENCES

- CAMERON, J. S., & WENDT, G. R. Chemical studies of behavior: VIII. Some anomalous onset effect of benzquinamide, a "tranquilizer." *Journal of Psychology*, 1964, 58, 265-275.
- GALBRAITH, G. G., & MOSHER, D. L. Associative sexual responses in relation to sexual arousal, guilt, and external approval contingencies. *Journal of Personality and Social Psychology*, 1968, 10, 142-147.
- HAEFNER, D. P. *Some effects of guilt-arousing and fear-arousing persuasive communications on opinion change*. (Tech. Rep. No. 1, Contract Nonr-668(12)) Washington, D. C.: Government Printing Office, 1956.
- KINSEY, A. C., POMEROY, W. B., MARTIN, C. E., & GEBHARD, P. H. *Sexual behavior in the human female*. Philadelphia: Saunders, 1953.
- LAZARUS, R. S., & ALFERT, E. The short-circuiting of threat by experimentally altering cognitive appraisals. *Journal of Abnormal and Social Psychology*, 1964, 69, 195-205.
- LAZARUS, R. S., SPEISMAN, J. C., MORDEKOFF, A. M., & DAVISON, L. A. A laboratory study of psychological stress produced by a motion picture film. *Psychological Monographs*, 1962, 76(34, Whole No. 553).
- MOSHER, D. L. The development and validation of a sentence completion measure of guilt. Unpublished doctoral dissertation, Ohio State University, 1961.
- MOSHER, D. L. The interaction of fear and guilt in inhibiting unacceptable behavior. *Journal of Consulting Psychology*, 1965, 29, 161-167.
- MOSHER, D. L. Measurement of guilt in females by self-report inventories. *Journal of Consulting and Clinical Psychology*, 1968, 32, 690-695.
- NOWLIS, V. Research with the Mood Adjective Checklist. In S. S. Thompkins & C. E. Izard (Eds.), *Affect, cognition, and personality*. New York: Springer, 1965.
- NOWLIS, V., & GREEN, R. F. *Factor analytic studies of mood*. (Tech. Rep. No. 1, Contract Nonr-668(12)) Washington, D. C.: Government Printing Office, 1964.
- NOWLIS, V., & NOWLIS, H. H. The analysis of mood. *New York Academy of Science*, 1956, 65, 345-355.
- OKEL, E., & MOSHER, D. L. Changes in affective states as a function of guilt over aggressive behavior. *Journal of Consulting Psychology*, 1968, 32, 265-270.
- ROBBINS, P. R. Self-reports of reactions to fear-arousing information. *Psychological Reports*, 1962, 11, 761-764.
- WICKENS, D. D., & MEYER, D. R. *Psychology*. (2nd ed.) New York: Dreyden, 1963.
- WILLINGHAM, C. *Eternal fire*. New York: Dell, 1964.

(Received August 15, 1968)

OBJECT-DISCRIMINATION LEARNING SET ACQUISITION IN YOUNG CEREBRAL PALSID CHILDREN IN RELATION TO TESTED AND RATED INTELLIGENCE¹

DOUGLAS PEARSON²

Pacific State Hospital

Object-discrimination learning set formation was explored as a potential means of assessing the intelligence of children with neuromotor handicaps. Forty learning set problems were administered to the following groups, each containing six Ss, of 2- to 3½-year-olds: cerebral palsied (CP) children who tested as retarded but who were rated as not being retarded; cerebral palsied children who tested and were rated as being retarded; nonhandicapped retardates; nonhandicapped children with average intelligence. The CPs who were rated as not retarded performed like the normals; both were significantly better learners than the other two groups. The discussion focused on the import of these findings for the future assessment of such children and the role that other variables play in such learning.

The high degree of objectivity, standardization, reliability, and validity of the Stanford Binet and the WISC are qualities largely responsible for their widespread use as intelligence tests for children. It is doubtful, however, whether these qualities are retained when these tests are given to children who, in some way, are grossly different from those on which the test was standardized. One such group of children would be those, like the cerebral palsied, who have motor and/or sensory handicaps. In many instances such children lack the requisite motor and/or sensory skills for doing a wide variety of test items. For example, they often do not have sufficient motor coordination to manipulate objects in a manner consistent with the purposes of testing. Likewise, many of these children are not able to talk with the degree of facility needed to express themselves on verbal test items. Finally, it might be argued that because of their handicaps, these children suffer some amount of environmental deprivation and therefore do not bring the same relevant experiential background to the

testing situation as do many of their nonhandicapped peers. As a result of this, it would appear that the examiner is often placed in the position of having to guess if such a child's failure to perform adequately on a test is indicative of an intellectual deficit, physical inability, or some combination of the two.

Attempts at circumventing this problem have been made by devising new tests (e.g., the Columbia Mental Maturity scale); by relying on pictorial recognition tests, such as the Peabody Picture Vocabulary Test, that place a minimal demand on the motor skills of the child (Dunn & Hartley, 1959); or by combining a potpourri of "usable" test items, modified where necessary, from a variety of existing intelligence tests for children (Haeussermann, 1952; Kogan, 1957; Taylor, 1959). These procedures have their drawbacks. In many instances they lack, to one degree or another, the objectivity, standardization, reliability, and validity of their more commonplace counterparts. In other instances, as with the use of pictorial recognition tests, the child's performance may be "sabotaged" by some of his more subtle problems such as a deficit in visual perception (Lacey, 1962) or less experience with his environment in comparison to most children his age.

In spite of these difficulties, the need for intellectual assessment early in the life of children with neuromuscular handicaps is

¹ This paper is based on a dissertation presented in partial fulfillment of the requirements for a PhD in Psychology at the University of California, Los Angeles. The author is grateful for the guidance and support of Wendell E. Jeffrey, major professor.

² Requests for reprints should be sent to Douglas Pearson, who is now at the University of Calgary, Department of Psychology, Calgary 44, Alberta, Canada.

often imperative. For one thing, there is always the concern on the part of the physician and the parent as to whether damage to the brain has resulted in impaired intellectual functioning. A more pressing reason for assessment in this case is that measured intelligence is an important factor in formulating future educational and social plans for such children. Here the gamut of choices is wide; they lie between placement in an institution for the retarded and the opportunity of staying at home and attending a regular public school. Obviously such decisions have considerable moment for both the parents and the child and deserve to be made on the most accurate information that can be provided.

In this investigation, the ability of cerebral palsied children to form object-discrimination learning sets was explored as an alternative to intelligence tests for assessing, on a gross level, the intellectual ability of such children. The primary reasons for adopting this alternative were that (a) the ability to form object-discrimination learning sets has already been shown to discriminate between groups of nonhandicapped children of varying levels of intelligence (Ellis, Girardeau, & Pryer, 1962; Girardeau, 1959; Kaufman & Peterson, 1958; Stevenson & Swartz, 1958); (b) the ability to perform on the learning-set task does not require speech or any but the barest amount of motor skill and, because of this, seemed like a technique that could readily be used with children with motor handicaps; and (c) it was felt that a measure of this kind of learning might ultimately prove to be a useful criterion for predicting how well such children might perform in a school setting.

There is a vast amount of experimental literature on the use of object-discrimination learning sets in this general context. Comparative studies, using this paradigm, have shown that the ability for such learning is a function of the rank of a species on the phylogenetic ladder. That is, a species ability to form learning sets will be superior to a species of lower phylogenetic rank, but inferior to a species of higher phylogenetic rank. Thus far, the rat, the cat, the marmoset, the raccoon, the spider monkey, and the rhesus monkey have been tested on object-discrimination

TABLE 1
DESCRIPTIVE STATISTICS OF THE FOUR GROUPS

Group	N	CA		IQ	
		M	SD	M	SD
Normals	6	32.7	4.6	103.5	5.4
CP "rated-not-retarded"	6	32.3	4.0	67.0	4.1
CP "rated-retarded"	6	32.7	6.0	60.0	6.6
Nonhandicapped retardates	6	31.0	4.4	61.0	9.6

Note.—CA is presented in months.

learning sets and, in general, their ability for such learning has ranked them in accordance with their position on the phylogenetic scale (Cotterman, Meyer, & Wickens, 1956; Harlow, 1958; Koronakos & Arnold, 1957; Miles, 1957; Miles & Meyer, 1956; Myers, McQuiston, & Miles, 1962; Schrier, 1966; Shell & Riopelle, 1957; Warren & Barron, 1956; Wright, Kay, & Sime, 1963).

Recent investigations with children below the age of seven have, as has already been mentioned, generally shown that the ability to form such sets is strongly related to classifications of measured intelligence (i.e., average, moderately retarded, etc.). Results have typically shown that normals reach a higher level of performance and take fewer trials to do so than retarded children (Ellis, 1958; Ellis, Girardeau, & Pryer, 1962; Girardeau, 1959; Harter, 1965; Stevenson & Swartz, 1958).

METHOD

Subjects

Four groups of six children, between 2 and 3½ years of age, were compared on their ability to form object-discrimination learning sets. Three of these groups were matched on IQ on the basis of their scores on the Cattell Infant Intelligence Scale; they contained children whose measured intelligence fell in the range of moderate to mild mental retardation (i.e., IQs between 50 and 75). Table 1 presents the descriptive statistics of the four groups. One of these three groups had retarded children who did not have any outstanding physical handicaps. The other two groups were comprised of cerebral palsied (CP) children whose physical handicaps quite conceivably interfered with their performance on a test like the Cattell or the Binet. One of these CP groups had children who were rated (see below) by nursery

school teachers and physical and occupational therapists (paramedical personnel very conversant with the children and how they functioned in a nursery school setting) as being retarded (the CP "rated-retarded" group), while the other group had children who were rated as not being retarded (the CP "rated-not-retarded" group). The fourth group was made up of nonhandicapped children with IQs in the average range of intelligence (i.e., between 95 and 110).

Rating Scales

Inasmuch as the major purpose of this experiment was to explore the use of learning sets as a gross measure of intellectual capacity that would be minimally influenced by the motor disabilities of CP children, it was necessary to have a second estimate of this ability that would be independent of both their handicap and their intelligence test scores. Such an estimate could then serve as a validity criterion against which to evaluate the performance of the CP Ss on the learning-set task. To this end a set of six rating scales was devised to be used by nursery school teachers and physical and occupational therapists. These scales were developed on the basis of the nursery school education literature (e.g., Read, 1960) pertaining to the kinds of age-related behaviors nursery school teachers might use in evaluating the intellectual and social development of their pupils. Five different age-related behavioral dimensions of this sort were developed into preliminary scales. They were: "Use of Play Material," "Social Interaction," "Expressive Communication," "Receptive Language," and "Social Awareness." A sixth scale asked the raters to rate each child as to his gross level of intelligence (i.e., average, borderline, severely retarded, etc.). Each scale had five points representing the range of age-related competency in the skill. The midpoint of the scale represented those sets of behaviors expected of most children around the age of 2½ to 3 years. The scale points were given descriptive labels. For example, the points on the "Expressive Communication" scale were (1) undifferentiated, (2) attention getting, (3) basic need communication, (4) social communication, and (5) interpersonal and fantasy communication. In addition, each of these points was anchored by one or more behavioral referents which, where possible, did not rely on the motor skills of the child. As an example of this, the referents for the points on the "Expressive Communication" scale labeled "social communication" were "in some manner communicates social needs and desires" and "expresses more subtle emotions like jealousy, shame."

These scales, in a preliminary form, were submitted to three highly experienced nursery school teachers for their suggested modifications. After the recommended changes were made, the scales were pretested by having three raters of the type to be used in the study independently rate each of six cerebral palsied children (not used in the study). The resulting reliability coefficients were between .76

and .86 and were found to be significantly different from zero ($p < .01$). In addition, the cumulative ratings over the five dimensional scales for each child were checked to see how well they agreed with the rating of the child's intelligence. The rank correlation coefficient for this match was .98 and found to be significantly greater than zero ($p < .01$).

In this study each cerebral palsied S was independently rated by two or more raters on the six scales. Raters were asked to make their judgments on each scale separate from ratings on any of the other scales. Since each scale ran from 1 to 5, Ss could receive a cumulative score ranging from 6 to 30. Because a score of 17 was the midpoint in the range of possible scores, it was made the cutoff point for putting the CP S into the "rated-retarded" or the "rated-not-retarded" group. In this case, Ss having a cumulative score between 6 and 17 were placed in the "rated-retarded" group while those having scores falling between 18 and 30 were placed in the "rated-not-retarded" group.

Apparatus

A Wisconsin General Test Apparatus (WGTA) was used to present the stimulus pairs to Ss. Since it had to be highly portable, the WGTA was made out of light material (1-inch-thick plywood) and was small enough to be easily transported and to fit on the top of any small table such as might ordinarily be found in a nursery school setting.

The experimental stimuli were 100 commonplace objects, like small toys or eating utensils, and solid geometric blocks of various shapes, colors, and sizes. It was felt that the use of such objects would make the stimuli in each pair more easily discriminated from one another since they were likely to differ on a number of dimensions such as size, shape, color, texture, and "meaning." In all, there were 100 such objects available for use in the study. Each object was tagged with a number between 1 and 100. The 40 test pairs used in the experiment were obtained by drawing, in succession, 40 pairs of numbers between 1 and 100, from a table of random numbers with the restriction that no number could be drawn more than once. In each of these 40 instances the objects corresponding to the numbers drawn formed a test pair.

Experimental Procedure

The initial phase of the procedure was given over to adapting S to the experimental situation and shaping him into making the required response. In general, shaping was simply directed at getting S to recognize that he could obtain a bit of reinforcement (sugar-coated cereal) by displacing, in some manner, one of two objects covering the foodwell on the discriminanda tray. Typically Ss needed only a few trials before they were able to make the desired response.

At this juncture the learning-set procedure was begun. Forty object-discrimination learning-set problems were presented to each S using a fixed-trial

regime (six trials per problem) with the left-right position of the reward object being varied from trial to trial according to a prearranged random sequence. On each trial of a given problem *S* was required to choose either by object displacement, head nod, or verbal identification the object which he thought had the reward under it. A noncorrectional procedure was used throughout the presentation of the test problems.

Primarily because of school schedules, 22 out of the 24 *Ss* were run on the experimental procedure on alternate days or three times a week. Some variation occurred here because of illness. Typically *Ss* were able to do about five learning-set problems a session so that it took an average of 3 weeks to complete the learning task. Two of the *Ss* were run on a daily schedule because they were available as *Ss* for a limited amount of time. In any event, it was felt that variations in scheduling experimental sessions was not critical because previous research on humans (Wischner, Braun, & Patton, 1962) had shown that various levels of skill in forming learning sets are retained over a period as long as 6 months.

The order of problem presentation, the trial sequence of reward-object position, and the designation of the reward object were randomized independently for each *S*. This was done to control for factors like position and object preferences or the inevitable variation in the difficulty level of the test pairs which could affect the experimental findings.

RESULTS

The basic response measure in this study was the number of correct responses on Trials 2 through 6 over consecutive blocks of eight learning-set problems. In all, there were five such blocks of problems. Table 2 presents the mean number of correct responses on each problem block for the four comparison groups (Figure 1 depicts the same data in graphic form). A Group (*G*) \times Problem Block (*PB*) repeated-measures analysis of variance was done on these mean scores to test for difference between the groups on the degree to and the rate at which they formed object-discrimination learning sets.

Amount Learned

The finding of a significant *F* ($F = 26.27$, $df = 3/20$, $p < .001$) for the group effect on the analysis of variance indicates that the groups did differ in the degree to which they formed object-discrimination learning sets. To identify which groups differed in this regard, the group means on each problem block were analyzed using a Duncan's multiple-range test. The results ($p = .05$: $R_1 = 7.46$, $R_2 =$

TABLE 2
MEAN NUMBER OF CORRECT RESPONSES ON
EACH PROBLEM BLOCK FOR THE FOUR
COMPARISON GROUPS

Group	Problem block				
	1	2	3	4	5
Normals	28.50	34.33	33.33	35.00	35.35
CP "rated-not-retarded"	28.00	27.57	29.83	32.33	33.33
CP "rated-retarded"	21.00	24.33	21.00	23.67	23.00
Nonhandicapped retardates	20.00	21.00	19.67	23.50	20.17

7.84, $R_3 = 8.07$) showed the mean performance of the CP "rated-not-retarded" group to be on a par with the normals by the end of the first block. Furthermore, both of these groups were found to perform at a level significantly above that of the CP "rated-retarded" and the nonhandicapped retarded groups on Blocks 3, 4, and 5. None of the other comparisons were found to be reliable.

Examination of the learning curves for the four groups (Figure 1) suggested that the normals performed at a much higher level sooner in the learning-set procedure than any of the other three groups. Hence, there was the possibility that their ability, in this regard, was primarily responsible for the difference found between the high and the low groups. Because it was critical to know if the CP "rated-not-retarded" group learned more than either of the low groups, a *G* \times *PB* analysis of variance and the Duncan multiple-range test were performed on the same data except that the scores of the normals were excluded. Results of both the analysis of variance ($F = 31.19$, $df = 2/20$, $p < .001$) and the Duncan, on Blocks 3 through 5 ($p = .05$: $R_1 = 7.46$, $R_2 = 7.84$), revealed that the high CPs had indeed learned more than either of the two low groups.

Finally, examination of the learning curves for each of the four groups (Figure 1) aroused some curiosity as to whether the two low groups, the CP "rated-retarded" and the nonhandicapped retardates, had formed learning sets to any significant degree. The learning

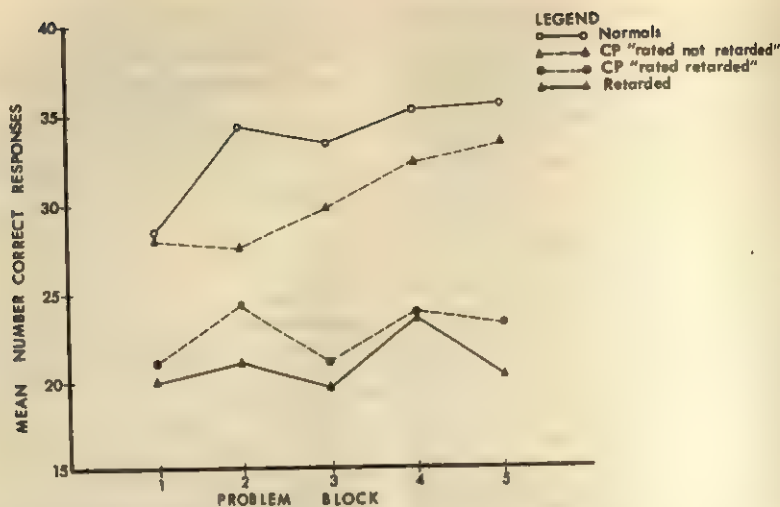


FIG. 1. Mean percentage of correct responses per problem block for each group.

trend for each group was subjected to a trend analysis using an analysis of variance technique. In this case, only a significant F for the linear component would mean that some learning had taken place across problem blocks. The results of these tests showed significant F s for the linear component on the group variable for the normals ($F = 7.65$, $df = 1/80$, $p < .01$), the CP "rated-not-retarded" ($F = 9.00$, $df = 1/80$, $p < .01$), and the CP "rated-retarded" ($F = 4.07$, $df = 1/80$, $p < .05$). Thus, only the nonhandicapped retardates, it turns out, failed to show any significant degree of learning on the experimental task.

Rate of Learning

The significant Groups \times Problems interaction ($F = 1.96$, $df = 12/80$, $p < .05$) indicates that there was indeed a difference between the groups in terms of their rates of learning. Several trend analyses using an analysis of variance technique were performed in order to further clarify the nature of this difference. A comparison of the learning trends for the two high groups, the normals, and the CP "rated-not-retarded," showed a significant linear component for the Groups \times Problems interaction ($F = 4.36$, $df = 1/40$, $p < .05$). On the basis of this finding, plus inspection of the learning curves for these two groups (Figure 1), it was felt that the

normals formed sets at a faster rate than the CP "rated-not-retarded," in the early problems, but that their rates became similar in the latter stages of learning. Another trend analysis showed that the same interaction effect ($F = 6.80$, $df = 1/80$, $p < .01$) was significant when tested on the two CP groups. Again, this finding, combined with inspection of the group learning curves (Figure 1) involved in the comparison, indicated that the CP "rated-not-retarded" learned at a faster rate than did the CP "rated-retarded." A comparison of the trends for the two low groups did not produce any significant findings.

DISCUSSION

The results of this study supported the hypothesis that the CP "rated-not-retarded" group's ability to form object-discrimination learning sets would be on a par with nonhandicapped children with average IQs and significantly better than the CP "rated-retarded" and a group of nonhandicapped retardates. The differences occurring between groups, especially the ones having nonhandicapped children, demonstrate once again the capacity for such learning is apparently related, in some positive fashion, to MA or IQ. The fact that the nonhandicapped retardates failed to show any degree of learning differs somewhat from what has been reported

on children at this IQ level. A probable reason for this is that the children in this group were run on 40 problems; this is far fewer than the 60-80 problems run on similar samples of children who have demonstrated some capacity for forming learning sets (e.g., Ellis, 1958; Girardeau, 1959; Kaufman & Peterson, 1958; Stevenson & Swartz, 1958).

Two other factors, task satiation and deficits in attention, were seen to play a disruptive role in the learning-set performance of a few Ss in each of the four groups. The first of these was most likely due to the fact that it took an average of 3-4 weeks to run an S through the learning-set procedure. During this time Ss seemed to hit a period or two when they became tired of the experimental task. When this occurred, these Ss no longer evidenced any concern for securing the reinforcement and they would then lapse into rigid or haphazard response patterns. With regard to the second fact, it was noted that some of the Ss would frequently not look at where they were reaching when they made their choice response. This was often the case with the nonhandicapped retardates, but it was an especially prevalent feature in the performance of the CP children. It was not possible to ascertain if this deficit was the result of the organic problems of some of the Ss or if some of the children, for whatever reason, had not yet learned to attend to environmental cues for any length of time in a problem-solving situation. The net result of this was twofold. One was an increase in individual variability which, because it was unsystematic in character, inflated the error variance. This would seem to imply that the trends noted in this study would have been even more statistically pronounced than they were if there was not this additional source of error. The other was a decrease in the reliability of the learning-set scores. Hence, it would not be possible to define a reliable and valid relationship between the learning-set scores and measured intelligence on the basis of these findings. Yet, the problem of using the learning-set procedure in this context does not appear to be insurmountable. Since the data showed that definite differences occurred between groups by the third problem block,

it is possible, for example, that fewer problems could have been run, thereby cutting down on some of the influence of the factors mentioned in the above. But in the broader perspective, the fact that variables of this nature influence such learning is by no means a major threat to clarifying the role of measurable aspects of intelligence in learning-set performance. Harter (1967) has already shown that it was possible to define the relationship between various kinds of incentive conditions, measured intelligence (both MA and IQ), and object-discrimination learning-set formation. It should be possible in future investigations to expand the knowledge as to the effect of variables like these in order to further clarify the contribution of MA and/or IQ to such learning.

One overriding result of this study was that the ability to form object-discrimination learning sets could be measured adequately on children whose handicaps possibly, for reasons outlined already, interfere with getting an adequate measure of intelligence. Perhaps the greatest importance should be attached to the fact that the CP "rated-not-retarded" group performed in a manner similar to an age-matched group of normals. It would appear, then, that the learning-set procedure allows such children to express abilities that are not tapped, and most certainly not predicted, by the IQ scores. Given the consistency with which children of average intellectual ability outperform retardates on this task, this finding lends credence to the possibility that current methods of IQ testing underestimate the ability of some children with motor and/or sensory handicaps. If future research shows that there is a relationship between learning-set formation and measured intelligence that is valid and reliable, the capacity for such learning might serve as one alternative way to assess the intellectual or educational potential of such children which would be more objective than the means now at our disposal.

REFERENCES

- COTTERMAN, T. E., MEYER, D. R., & WICKENS, D. O.
Discrimination reversal learning in marmosets.
Journal of Comparative and Physiological Psychology, 1956, 49, 539-541.

- DUNN, L. M., & HARTLEY, R. K. Comparability of the Peabody, Ammons, Van Alstyne, and Columbia test scores with cerebral palsied children. *Exceptional Children*, 1959, 26, 70-74.
- ELLIS, N. R. Object-quality learning sets in mental defectives. *Journal of Comparative and Physiological Psychology*, 1958, 51, 79-81.
- ELLIS, N. R., GIRARDEAU, F., & PRYER, M. W. Analysis of learning sets in normal and severely defective humans. *Journal of Comparative and Physiological Psychology*, 1962, 55, 556-570.
- GIRARDEAU, F. The formation of learning sets in mongoloid and normal children. *Journal of Comparative and Physiological Psychology*, 1959, 52, 566-570.
- HAEUSSERMANN, E. Evaluating the developmental level of cerebral palsied children. *Journal of Genetic Psychology*, 1952, 80, 3-28.
- HARLOW, H. F. The evolution of learning. In A. Roe & G. Simpson (Eds.), *Behavior and evolution*. New Haven: Yale University Press, 1958.
- HARTER, S. Discrimination learning sets as a function of IQ and MA. *Journal of Experimental Child Psychology*, 1965, 2, 31-43.
- HARTER, S. Mental age and motivational factors in the discrimination learning set performance of normal and retarded children. *Journal of Experimental Child Psychology*, 1967, 5, 123-141.
- KAUFMAN, M. E., & PETERSON, W. H. Acquisition of a learning set by normal and mentally retarded children. *Journal of Comparative and Physiological Psychology*, 1958, 51, 619-621.
- KOGAN, K. L. A method of assessing intellectual capacity in preschool cerebral palsied children. *Journal of Clinical Psychology*, 1957, 13, 54-56.
- KORONAKOS, C., & ARNOLD, W. J. The formation of learning sets in rats. *Journal of Comparative and Physiological Psychology*, 1957, 50, 11-14.
- LACEY, H. M. Pre-conditions for the psychological evaluation of the young cerebral palsied child. *Cerebral Palsy Review*, 1962, 23, 13-14.
- MEYERS, W. J., MCQUISTON, M. D., & MILES, R. C. Delayed-response and learning-set performance in cats. *Journal of Comparative and Physiological Psychology*, 1962, 55, 515-517.
- MILES, R. C. Learning set formation in the squirrel monkey. *Journal of Comparative and Physiological Psychology*, 1957, 50, 356-357.
- MILES, R. C., & MEYER, D. R. Learning sets in marmosets. *Journal of Comparative and Physiological Psychology*, 1956, 49, 219-222.
- READ, K. H. *The nursery school: A human relation laboratory*. Philadelphia: Saunders, 1960.
- SCHRIER, A. M. Learning set formation in three species of macaque monkeys. *Journal of Comparative and Physiological Psychology*, 1966, 61, 866-868.
- SHELL, W. F., & RIOPELLE, A. J. Discrimination learning in raccoons. *Journal of Comparative and Physiological Psychology*, 1957, 50, 585-587.
- STEVENSON, W. H., & SWARTZ, J. D. Learning set formation in children as a function of IQ level. *Journal of Comparative and Physiological Psychology*, 1958, 51, 755-757.
- TAYLOR, E. M. *Psychological appraisal of children with cerebral palsy*. Cambridge: Harvard University Press, 1959.
- WARREN, J. M., & BARRON, A. The formation of learning sets in cats. *Journal of Comparative and Physiological Psychology*, 1956, 49, 227-231.
- WISCHNER, G. J., BRAUN, H. W., & PATTON, R. A. Acquisition and long-term retention of an object quality learning set by retarded children. *Journal of Comparative and Physiological Psychology*, 1962, 55, 518-523.
- WRIGHT, P. I., KAY, H., & SIME, M. E. The establishment of learning sets in rats. *Journal of Comparative and Physiological Psychology*, 1963, 56, 200-203.

(Received August 16, 1968)

INTERNAL-EXTERNAL CONTROL DIMENSION AS RELATED TO ACCIDENT AND SUICIDE PRONENESS

CHRISTENE BLANTON WILLIAMS¹ AND JAMES B. NICKELS²

University of Manitoba

A 239-item inventory consisting of Rotter's Internal-External scale, Keehn's Accident Index, Farberow and Devries' MMPI Suicide scale, and Devries' Potential Suicide Personality Inventory was administered to 235 introductory psychology students. As predicted, externally oriented Ss generally scored higher on the suicide potentiality scales than internally oriented Ss (with the two suicide scales correlating highly with each other). Contrary to predictions, scores on the accident proneness scale correlated positively with scores on both the externality and suicide potentiality scales. Possible explanations for these results were discussed.

A review of the literature on accident and suicide proneness, on the one hand, and one's perception of reinforcement as internally or externally controlled on the other hand, suggests that these areas may be related. That is, there is reason to believe that accident-prone individuals may be internally oriented and suicide-prone individuals may be externally oriented.

The internal-external control dimension is based on Rotter's (1954) social learning theory and describes two quite different world views. The internally oriented individual believes that reinforcements are contingent on certain aspects of his own behavior, such as competence and skill; while, conversely, the externally oriented individual believes that such reinforcements are determined by forces independent of his own behavior, such as fate, chance, luck, or other individuals.

In recent years there have been a number of studies showing differences in behavior as a function of the locus-of-control dimension (Lefcourt, 1966b; Rotter, 1966). Several investigators have demonstrated that the acquisition and extinction of expectancies depends on whether Ss perceive a task as skill- or chance-determined (James & Rotter, 1958; Nickels & Tolen, 1968; Phares, 1957; Rotter, Liverant, & Crowne, 1961). In fact, a scale has been developed to measure individual

differences in generalized locus-of-control expectancy (Rotter, 1966). Through the use of this scale and other procedures, data have been obtained which indicate that internal scorers differ from external scorers in a variety of ways. For example, they differ in preferences for skill versus chance activities (Schneider, 1968), perceptual threshold variation (Phares, 1962), delay in decision-making (Rotter & Murray, 1965), latent learning performance (Getter, 1966), unusual shifts in expectancy (Battle & Rotter, 1963; James, 1957), memory for various kinds of information (Seeman, 1963), tendencies to forget failure experiences (Efran, 1963), degrees of conformity (Crowne & Liverant, 1963), resistance to subtle influence (Gore, 1962), attempts to control the environment (Gore & Rotter, 1963; Strickland, 1965), risk-taking behavior (Liverant & Scodel, 1960), attitudes toward information and social influence (Seeman & Evans, 1962), seeking of relevant information (Davis & Phares, 1967), and achievement (Butterfield, 1964; Crandall, Katkovsky, & Crandall, 1965).

More recently Hersch and Scheibe (1967) treated locus-of-control as a personality dimension which they related to items and scales of the California Psychological Inventory (Gough, 1964) and the Adjective Check List (Gough & Heilbrun, 1965). Results indicated that internal scorers as compared to external scorers were higher on the Dominance, Tolerance, Capacity for Status, Good Impression, Social Presence, Sociability, Psychological Mindedness, Intellectual Efficiency,

¹ Requests for reprints should be sent to Christene Blanton Williams, Department of Psychology, University of Manitoba, Winnipeg, Canada.

² The authors wish to express their gratitude to Florence Vantress for her assistance in data analysis.

Self-Control, Self-Acceptance, Well-Being, Responsibility, Achievement via Conformance, and Achievement via Independence scales of the California Psychological Inventory. Also, internal scorers were found to be higher on the Adjective Check List scales of Defensive-ness, Self-Confidence, Achievement, Dominance, Endurance, and Order, but lower on scales of Succorance and Abasement. More specific analysis of individual items revealed that internal scorers viewed themselves as clever, efficient, egotistical, enthusiastic, independent, self-confident, ambitious, assertive, boastful, conceited, conscientious, deliberate, persevering, clear-thinking, dependable, determined, hardheaded, industrious, ingenious, insightful, organized, reasonable, and stubborn (as determined by these adjectives being checked significantly more often by internal Ss). According to the researchers, internally oriented persons were more likely to describe themselves as active, striving, achieving, powerful, independent, and effective. On the other hand, external scorers checked only one adjective significantly often—that of “self-pitying” (Hersch & Scheibe, 1967). However, some evidence suggested that external scorers were more anxious and maladjusted than internal scorers. Supporting this view, Feather (1967) attributed greater anxiety and neuroticism to external scorers.

Investigations relevant to (but not previously related to) the locus-of-control variable include studies on accident proneness and suicide proneness. The accident-prone individual was described by Jenkins (1956) as a talkative, rule-breaking, independent, self-confident, unrestrained, nonconforming, attention-seeking, brazen, and reckless individual who was cooperative only when it benefited him. As compared to “accident nonrepeaters,” they displayed greater superiority, boldness, initiative, and leadership, but were also more distractible, aggressively outspoken, impolitic in frankness, and grandiose in thinking. In addition, they felt less sensitive, less guilty, and less ashamed of themselves than non-repeaters, and had their feelings hurt less often by others. (In fact, they seemed to obtain satisfaction or thrill in minor injuries, in being in pain, or in being hurt by others.) When frustrated, however, accident-prone

persons had less control of their feelings and so became quite upset and angry. They also revealed a greater interest in disciplinary and power vocations and less signs of emotional maturity and mental adequacy. Other characteristics attributed to accident-prone individuals portray them as being extravertive (Fine, 1963; Keehn, 1961), imprudent (Fordyce, 1964), daring and adventurous (Kunce & Brewer, 1966; Kunce & Worley, 1966), and quick, decisive, impulsive, and independent (Alexander, 1949; Dunbar, 1943). Interestingly, one group of researchers (Kunce & Brewer, 1966) suggested that accident proneness may be incompatible with chronic and severe maladjustment, whereas another group (Selzer, Payne, Gifford, & Kelly, 1963) found a high incidence of psychiatric illness among accident-prone persons.

On the basis of previous studies, a comparable picture of the suicide-prone (potentially suicidal) individual can be drawn. He has been characterized as rigid (Fairbank, 1932; Vinoda, 1966; Wall, 1944), introversive and asocial (Hendin, 1950; Hopkins, 1937; Williams, 1936), timid and dependent (Batchelor & Napier, 1954; Faris, 1934), easily embarrassed, overly sensitive, restrained in forming friendships, and lacking in self-confidence and initiative (Farberow & Devries, 1967), feeling fearful, worthless, and unwanted (Devries, 1966), irritable and resentful (Lester, 1967; Stoneman & Perth, 1935), and poorly integrated, weak in character, or abnormal in personality (Andics, 1947; Raphael, Power, & Berridge, 1937; Sainsbury, 1955; Schneider, 1954—reviewed by Robin, 1956; Siewers & Davidoff, 1942). In a study contrasting suicidal and nonsuicidal patients with respiratory or cardiac illnesses (Farberow, McKelligott, Cohen, & Darbonne, 1966), the former group was found to be more dependent, impulsive, alert, anxious, agitated, apprehensive, depressed, and distressed over their illnesses. They tended to be “bad patients,” for they were also hostile and intolerant of pain and discomfort as well as complaining and demanding of special attention from hospital personnel. On the basis of an investigation of another patient population (Dean, Miskimins, DeCook, Wilson, & Maley, 1967), the researchers found suicidal patients to display

less overt anger, inappropriate behavior, and benefit from social activities (but more apathy) than nonsuicidal control patients. Interestingly, the suicidal patients in this study were labeled "cooperative" and rated higher than controls in "adjustment," while other researchers seemingly stereotyped them as "uncooperative" (Farberow et al., 1966) and lacking in "adjustment" (Andics, 1947; Batchelor & Napier, 1954; Devries, 1966). A parallel inconsistency may be pointed out in that Dean and his colleagues, as well as Winfield and Sparer (1953), found suicides to be less extrapunitive than controls, whereas Farberow (1950) found suicides to be more extrapunitive than controls, and both Fisher and Hinds (1951) and Vinoda (1966) found no differences between the attempted suicide population and psychiatric patients (except that suicidal persons were more rigid and revealed more general hostility).

Although previous investigators have not compared the locus-of-control variable with either accident or suicide proneness, several researchers have specifically related the latter two areas. For example, Litman and Tabachnick (1967) described the accident-prone person as active, counterphobic or counterdepressive, impulsive, quick, decisive, independent, adventurous, rebellious, and defiant of authority. On the other hand, these same researchers described the suicide-prone person as masochistic, depressive, dependent, passive, immobilized, constricted, and disinterested in physical activity. In other words, the former individual was viewed as a "risk-taking winner," the latter as a "danger-avoiding loser." Also, Preston (1964) found that accident-prone persons were more extrapunitive, less impunitive, and rather barrier-oriented (obstacle-dominance). Conversely, he found that suicide-prone persons were less extrapunitive, more impunitive, and rather ego-oriented (ego-defense). The results on extrapunitiveness tend to confirm the findings of several investigators (Dean et al., 1967; Winfield & Sparer, 1953) discussed above. In another study (Tabachnick, Litman, Osman, Jones, Cohn, Kasper, & Moffat, 1966), accident victims were found to be characteristically active, self-satisfied, and self-aggrandizing, but concerned with their appearance, whereas

suicide victims were dependent, self-critical, and self-punishing, but unconcerned with their appearance. Also, the former tended to be idealized by their survivors, whereas the latter tended to be depreciated by them. Although within a year prior to their deaths, accident victims had contemplated moving (or had actually moved) into situations of personal growth and responsibility, suicide victims just before their deaths had encountered the loss of a significant person or a feeling of failure (or being unloved). In addition, accident-prone persons were distantly integrated with and relatively autonomous of individuals in their environments, whereas suicide-prone persons tended to be closely integrated with and highly dependent on individuals in their environments. Finally, persons dying by accidents were rather exhibitionistic and frequently upset when slighted and criticized by others, whereas persons dying by suicide were rather depressive and frequently upset by the loss of a significant other or the deprivation of important resources. As an interesting aside, two other investigators cast doubts on these results by claiming that suicide-prone persons display exhibitionism (Teicher, 1947) and a relative insensitivity to "hurt feelings" (Jenkins, 1956).

From the above summary of research on the internality-externality and accident-suicide variables, it appears that many of the characteristics attributed to the internal orientation are those typically associated with accident proneness, and many of the characteristics attributed to the external orientation are those typically associated with suicide proneness. That is, both internally oriented and accident-prone persons tend to be independent, self-confident, brazen, boastful, enthusiastic, active, and achieving; whereas both externally oriented and suicide-prone persons tend to be dependent, guilty, anxious, timid, depressive, and unachieving. The former reveal their initiative, dominance, and power orientations, and a recent gain in personal growth and responsibility; the latter reveal their lack of initiative, their lack of self-control, and a recent loss of significant persons and resources in their lives. In addition, the former appear more self-acceptant, self-satisfied, and self-aggrandizing; the latter more

self-critical, self-punishing, and self-pitying. Finally, the former seem to be more sociable, more concerned with their appearance and making a good impression, and more idealized by significant others; the latter more asocial, more unconcerned with their appearance and making a good impression, and more depreciated by significant others.

A close approximation to the above description of the internality-accident combination is Litman and Tabachnick's (1967) statements:

the accident-prone individual is active. . . independent, adventurous, and likes excitement. . . . He cannot tolerate the feeling of being boxed in or losing autonomy [p. 253].

A major responsibility for fatal accidents must also be assigned to counterphobic and counterdepressive attitudes. . . . Some people try to avert depression by committing themselves to endeavors that exceed their capacities. This may lead to bungled actions which are potentially fatal. . . . Counterphobic persons gain self-esteem by taking drugs without becoming addicted, or by driving while tired, fighting fatigue, and struggling to stay awake. . . . The accident-prone individual may place himself in risky situations to strengthen his feelings of invulnerability and omnipotence . . . [pp. 256-257].

On the other hand, Weiss (1957) has clearly described the externality-suicide combination in his statement:

Many suicidal attempts have at least in part the character of a gamble with death, a sort of Russian roulette, the outcome of which depends to some extent on chance. The attempts are consciously or unconsciously arranged in such a manner that the lethal probability may vary from almost certain survival to almost certain death; and "fate"—or at least some force external to the conscious choice of the person—is compelled in some perhaps magical way to make the final decision [p. 21].

The purpose of the present study was to determine whether accident-prone persons view their reinforcements as contingent on their own skill, capability, proficiency, or expertise (internality), whereas suicide-prone persons view their reinforcements as contingent on fate, chance, luck, or other persons (externality). Accordingly, two hypotheses were advanced: First, measures of the external locus-of-control philosophy will correlate positively with measures of suicide proneness but negatively with measures of accident

proneness. Second, measures of suicide proneness will correlate positively with one another but negatively with measures of accident proneness.

METHOD

Subjects

The Ss were 235 introductory psychology students (114 males and 121 females) at the University of Manitoba, who voluntarily participated in this experiment as part of their course requirement.

Instruments

A single questionnaire was administered to groups of 40 to 60 Ss. The questionnaire consisted of all items comprising four separate scales, namely, the Internal-External scale (I-E; Rotter, 1966); the Accident Index (AI; Keehn, 1959); the MMPI Suicide scale (MMPIS; Farberow & Devries, 1967); and the Potential Suicide Personality Inventory (PSPI; Devries, 1966). The I-E scale was kept intact, but the items from the other instruments were placed into the final questionnaire through the use of a table of random numbers.

The I-E is a 23-item (plus 6 fillers) forced-choice instrument reduced from an original pool of 100 items developed on a priori grounds. For each non-filler item an internal belief is compared to an external belief, and extensive normative data have been obtained on various populations under various circumstances.

The AI is a 38-item acquiescence-type instrument reduced from an original pool of 41 items developed as a listing of various accidents and minor mishaps. One hundred (mostly Arabic) male university students between 18 and 25 years of age served as the normative population.

The MMPIS is a 52-item true-false scale extracted from the short form of the MMPI. Normative data were obtained on male neuropsychiatric hospital patients separated into (a) original groups of 21 committed suicides, 32 attempted suicides, 32 threatened suicides, and 32 nonsuicidal patients; (b) replication groups of 22 committed suicides, 47 attempted suicides, 50 threatened suicides, and 48 nonsuicidal patients; and (c) cross-validation groups of 50 threatened suicides and 50 nonsuicidal patients. The median ages of the various groups ranged between 32 and 37.

The PSPI is a 50-item (plus 5 fillers) acquiescence-type instrument reduced from 333 critical incidents relevant to suicide. The original items were obtained from research articles on suicide. Normative data were obtained on male neuropsychiatric hospital patients separated into groups of 33 attempted suicides, 55 threatened suicides, 42 attempted and threatened suicides, and 83 nonsuicidal patients. Fifty-two percent of the suicidal patients were between 17 and 39 years of age, whereas 45% of the nonsuicidal patients fell within this age range.

TABLE 1

MEANS AND STANDARD DEVIATIONS FOR MEASURES OF
EXTERNALITY, ACCIDENT PRONENESS, AND
SUICIDE POTENTIALITY

Measure	<i>M</i>		<i>SD</i>	
	Males	Females	Males	Females
Externality (I-E)	9.28	9.41	4.28	2.36
Accident proneness (AI)	14.62	15.91	4.58	5.30
Suicidal potentiality (MMPIS)	14.91	16.12	8.40	8.06
Suicidal potentiality (PSPI)	23.03	22.57	3.34	3.30

Note.—For males, $N = 114$; for females, $N = 121$.

Procedure

Intercorrelations among the measures of externality, accident proneness, and suicide potentiality were based on the Pearson product-moment correlation and were computed separately for each sex. A one-way analysis of variance on extreme I-E scores was also computed in order to provide a further test of the relationship between externality on the one hand and AI, MMPIS, and PSPI scores on the other.

RESULTS

Means and standard deviations for measures of externality, accident proneness, and suicide potentiality were computed separately for each sex and appear in Table 1.

Intercorrelations among these same measures appear in Table 2.

All but two correlations reached statistical significance, with three being highly significant. As predicted in the first hypothesis, externality was directly related to suicide

potentiality as measured by the MMPIS (for both sexes) and the PSPI (for females only). In opposition to this same hypothesis, externality was neither directly related to male PSPI scores nor inversely related to AI scores (for either sex). The second hypothesis was also partially confirmed, since correlations between the two suicide potentiality scales (MMPIS \times PSPI) were highly significant in a positive direction. Contrary to this hypothesis, however, AI scores did not correlate negatively with MMPIS or PSPI scores. In fact, all correlations obtained in the present study were positive, and the third highest coefficient was that resulting from the correlation of AI and female MMPIS scores.

The one-way analysis of variance revealed that the 114 "clearly" external Ss (those 54 males and 60 females with externality scores greater than 9) had significantly higher accident- and suicide-proneness scores than the 65 "clearly" internal Ss (those 31 males and 34 females with externality scores less than 7), regardless of whether the scores were based on the AI ($F = 12.80$, $p < .001$), the MMPIS ($F = 20.53$, $p < .001$), or the PSPI ($F = 10.53$, $p < .01$). In other words, through the use of the extreme I-E cutoff points suggested by Eisenman and Platt (1968) from Rotter's normative data (1966), the researchers obtained additional confirmation for the results presented in Table 2.

DISCUSSION

The results of the present study, taken as a whole, suggest that externality, accident

TABLE 2

INTERCORRELATIONS AMONG MEASURES OF EXTERNALITY, ACCIDENT PRONENESS,
AND SUICIDE POTENTIALITY

Measure	Externality (I-E)		Accident proneness (AI)		Suicide potentiality (MMPIS)	
	Males	Females	Males	Females	Males	Females
Accident proneness (AI)	.20*	.14	—	—	—	—
Suicide potentiality (MMPIS)	.25**	.18*	.29**	.42***	—	—
Suicide potentiality (PSPI)	.14	.18*	.24*	.27**	.56***	.47***

Note.—For males, $N = 114$; for females, $N = 121$.

* $p < .05$, two-tailed.

** $p < .01$, two-tailed.

*** $p < .001$, two-tailed.

prone, and suicide potentiality are personality traits which vary together in the college population. The unexpected finding that AI scores correlate directly with measures of both externality and suicide proneness—rather than inversely as had been predicted—may reflect the lack of validity of the AI. Although the AI was developed on a university student norm group, it may not validly measure accident proneness either for present Ss (since most were Canadian rather than Arabic) or for any but the most honest accident repeater (since the AI requests truthful answers to obvious questions about a rather “touchy” problem area). On the other hand, perhaps the lack of I-E differentiation between accident and suicide proneness in the present research reflects the dissimilarity between the male and female normal college sample studied herein and the exclusively male neuropsychiatric hospital samples used to establish MMPIS and PSPI norms. Finally, the randomizing of AI, MMPIS, and PSPI items may have contributed in some way to the unexpected commonality among all four personality measures.

Nevertheless the question can still be asked, “If the four personality measures are valid, how can the results of the present study be explained?” Perhaps the accident-prone Ss were internally oriented as originally hypothesized but circumstances inflated their externality. First, the environment in which Ss function may to some extent influence whether they appear internally or externally oriented. Internals have looked like externals when reinforcements seemed to be chance-determined (Rotter & Mulray, 1965), and externals have looked like internals when reinforcements seemed to be skill-determined (Lefcourt & Ladwig, 1965). According to Hersch and Scheibe (1967) an individual may describe himself as an external because he is in a highly competitive situation where the actions of others may indeed have great relevance for the success of his own efforts. Therefore, Ss used in the present study (college students) may have viewed their present academic reinforcements as contingent on others, and so may have responded as if they were externally oriented individuals,

even though their basic world views may have been internal.

Second, both accident proneness (Litman & Tabachnick, 1967, pp. 251–252) and suicide proneness (Shneidman, 1965, pp. 319–320) have been called transient personality characteristics rather than permanent character traits. Therefore, accident-prone persons in the present study could have been internally oriented during certain transition periods in their lives, at which times they reacted to stress by engaging in reckless and rule-breaking behavior resulting in accidents. Afterward, however, when the stress lessened (and they agreed to participate in the present study), they indicated both their previous, internally oriented accidents but their present locus-of-control philosophies, which for a large number of Ss could have changed to an external orientation.

Third, there is evidence that accident-prone persons tend to “forget” their accidents (LeShan & Brame, 1953), and internals tend to “forget” their failure experiences (Efran, 1963). This “forgetfulness” could be a factor if the internal individual who views himself as efficient, competent, deliberate, and clear thinking fails to maintain this image by having accidents (e.g., crashing cars, cutting himself, bumping into things, etc.). He might tend to forget these failures and therefore answer the AI as if he were really nonaccident prone. Such behavior on a large scale could reverse the relationship between internality and accident proneness, and could account for present results.

Looking at the results from the reverse standpoint, perhaps the accident-prone Ss in the present study were, in fact, externally oriented. The problem then becomes one of explaining how the characteristics of accident proneness can seem to be consistent with the internal orientation and yet not be upon closer scrutiny. First, accident-prone persons may resemble internally oriented individuals in many respects (e.g., independence and self-aggrandizement), even though they may have certain other qualities—more essential to the concept of accident proneness—which internally oriented persons typically lack (e.g., recklessness and rule-breaking).

Second, although the outward appearance of accident-prone persons may suggest the self-control and self-acceptance of internally oriented individuals, the inner dynamics of the case (Litman & Tabachnick, 1967) may reveal that their impressive demeanor is actually a superficial denial of the urge to regress, to be passive, and to engage in escape, guilt, and revenge fantasies. In other words, in most instances they may do a superb job of hiding their close kinship to the suicide-prone person, although the present study uncovered this relationship.

As a preliminary formulation, we suggest that the crucial factor which suicide prone and accident prone have in common is a weakness of the integrating and synthesizing functions of the ego which allows two distinct patterns of activity and passivity to coexist. In this sense, suicide-prone and accident-prone states are reciprocal [Litman & Tabachnick, 1967, p. 257].

Perhaps the various measures of accident and suicide proneness correlate because they tap this common factor of inner dynamics.

If this common factor is labeled "maladjustment," then measures of accident and suicide proneness might be expected to correlate positively with I-E scores, which themselves have been found to relate directly to measures of maladjustment (Cromwell, Rosenthal, Shakow, & Kahn, 1961; Feather, 1967; Hersch & Scheibe, 1967; Rotter, 1966). Furthermore, even though "general anxiety" may (Feather, 1967; Watson, 1967) or may not (Efran, 1963) correlate positively with measures of externality, one study (Tolor & Reznikoff, 1967) found that externals revealed a greater amount of overt "death anxiety" than internals. Such findings on maladjustment and death anxiety may well account for the positive correlations among measures of externality, accident proneness, and suicide potentiality.

Third, although in the present study most accident- and suicide-prone persons may be externally oriented, any one-to-one pairing of locus-of-control, accident-proneness, and suicide-potentiality variables probably represents a gross oversimplification. In contrast to the original hypotheses, some investigators have linked accident proneness to externality and suicide proneness to internality. For example, the relationship between accident

proneness and externality has been clearly made.

The ego of the accident prone refuses to accept responsibility for his actions. . . . At the same time, there is no feeling of ego-alien "driving" toward accidents. They simply "happen," and "fate" or "luck" seems to be against the person [LeShan, 1952, footnote, p. 77].

Conversely, researchers in the area of suicide prevention have clearly indicated that some persons play a conscious and deliberate role in their own demise. Whether such individuals are called implementers (Farberow, Shneidman, & Leonard, 1963) or psyde-initiators (Shneidman, 1963), they may certainly be described as internally oriented.

A Psyde-initiator believes that he will suffer cessation in the fairly near future—a matter of days or weeks—or he believes he is failing and, not wishing to accommodate himself to a new (and less effective and less virile) image of himself, does not wish to let "it" happen to him. Rather, *he* wants to play a role in its occurrence. Thus he will do it for himself, at his own time, and on his own terms. . . . We find, on occasion, a case in which an older person, hospitalized in a general medical hospital, in the terminal stages of a fatal disease will, with remarkable and totally unexpected energy and strength, take the tubes and needles out of himself, climb over the bed rails, lift a heavy window, and throw himself to the ground several stories below [Shneidman, 1963, p. 213].

To further complicate the picture, accidents may sometimes represent self-destructive intentions (MacDonald, 1964; Menninger, 1936) as well as attempts to solve personal problems in dangerous surroundings (Behan & Hirshfield, 1963). Similarly, so-called suicides may represent an unintentioned or accidental ending to one's life as well as an example of deliberate self-destruction (Shneidman, 1963).

As the present researchers view the complex interaction among the locus-of-control, accident-proneness, and suicide-potentiality variables, they hypothesize that, aside from a person's physical condition, his death may to a large extent depend on (a) his desire to live or die, (b) his expectancy that he will live or die (within a prescribed period), and (c) his locus-of-control philosophy. Future studies may well consider the degree to which the above factors can predict fatal and non-

fatal accidents and attempted and completed suicides. In addition, through the use of the locus-of-control variable, further attempts can be made to differentiate those persons or conditions ripe for both accidents and suicides from those persons or conditions ripe for one but not the other (Henry & Short, 1951; Selzer & Payne, 1962).

Further investigations of accident and suicide proneness may well reveal the close relationship between these variables and that of locus of control. Since Lefcourt (1966a) has already indicated the importance of the I-E dimension for work in psychopathology and psychotherapy, perhaps the present study can offer a slight glimpse of the benefits for extending locus-of-control studies into the taboo areas of accidents, suicides, and death.

REFERENCES

- ALEXANDER, F. The accident prone individual. *Public Health Reports*, 1949, **64**, 357-361.
- ANDICS, M. VON. *Suicide and the meaning of life*. London: Hodge, 1947.
- BATCHELOR, I. R. C., & NAPIER, M. B. The sequelae and short-term prognosis of attempted suicide. The results of a one year follow-up of 200 cases. *Journal of Neurological Psychiatry*, 1954, **17**, 261-266.
- BATTLE, E. S., & ROTTER, J. B. Children's feelings of personal control as related to social class and ethnic group. *Journal of Personality*, 1963, **31**, 482-490.
- BEHAN, R. C., & HIRSCHFELD, A. H. The accident process. *Journal of the American Medical Association*, 1963, **186**, 300-306.
- BUTTERFIELD, E. C. Locus of control, test anxiety, reactions to frustration, and achievement attitudes. *Journal of Personality*, 1964, **32**, 298-311.
- CRANDALL, V. C., KATKOVSKY, W., & CRANDALL, V. J. Children's beliefs in their own control of reinforcement in intellectual-academic achievement situations. *Child Development*, 1965, **36**, 91-109.
- CROMWELL, R. R., ROSENTHAL, D., SHAKOW, D., & KAEN, T. Reaction time, locus of control, choice behavior, and descriptions of parental behavior in schizophrenic and normal subjects. *Journal of Personality*, 1961, **29**, 363-380.
- CROWNE, D. P., & LIVERANT, S. Conformity under varying conditions of personal commitment. *Journal of Abnormal and Social Psychology*, 1963, **66**, 547-555.
- DAVIS, W. L., & PHARES, J. E. Internal-external control as a determinant of information seeking in a social influence situation. *Journal of Personality*, 1967, **36**, 547-561.
- DEAN, R. A., MISKIMINS, W., DECOOK, R., WILSON, L. T., & MALEX, R. F. Prediction of suicide in a psychiatric hospital. *Journal of Clinical Psychology*, 1967, **23**, 296-301.
- DEVRIES, A. G. A potential suicide personality inventory. *Psychological Reports*, 1966, **18**, 731-738.
- DUNBAR, F. *Psychosomatic diagnosis*. New York: Paul B. Hoeber, 1943.
- EFRAN, J. S. Some personality determinants of memory for success and failure. Unpublished doctoral dissertation, Ohio State University, 1963.
- EISENMAN, R., & PLATT, J. J. Birth order and sex differences in academic achievement and internal-external control. *Journal of General Psychology*, 1968, **78**, 279-285.
- FAIRBANK, R. E. Suicide: Possibilities of prevention by early recognition of some danger signals. *Journal of the American Medical Association*, 1932, **2**, 1711-1714.
- FARBEROW, N. L. Personality patterns of suicidal mental hospital patients. *Genetic Psychology Monographs*, 1950, **42**, 3-79.
- FARBEROW, N. L., & DEVRIES, A. G. An item differentiation analysis of MMPI's of suicidal neuropsychiatric hospital patients. *Psychological Reports*, 1967, **20**, 607-617.
- FARBEROW, N. L., MCKELLIGOTT, J. W., COHEN, S., & DARBONNE, A. Suicide among patients with cardiorespiratory illnesses. *Journal of the American Medical Association*, 1966, **195**, 128-134.
- FARBEROW, N. L., SHENEDMAN, E. S., & LEONARD, C. V. *Suicide among general medical and surgical hospital patients with malignant neoplasms*. (Medical Bulletin 9) Washington, D. C.: Veterans Administration, 1963.
- FARIS, R. E. L. Cultural isolation and the schizophrenic personality. *American Journal of Sociology*, 1934, **48**, 155.
- FEATHER, N. T. Some personality correlates of external control. *Australian Journal of Psychology*, 1967, **19**, 253-260.
- FINE, B. J. Introversion-extraversion and motor vehicle driver behavior. *Perceptual and Motor Skills*, 1963, **12**, 95-100.
- FISHER, S., & HINDS, E. The organization of hostility controls in various personality structures. *Genetic Psychology Monographs*, 1951, **44**, 3-68.
- FORDYCE, W. E. Personality characteristics in men with spinal cord injury as related to manner of onset of disability. *Archives of Physical Medicine and Rehabilitation*, 1964, **45**, 321-325.
- GETTER, H. A personality determinant of verbal conditioning. *Journal of Personality*, 1966, **34**, 397-405.
- GORE, P. M. Individual differences in the prediction of subject compliance to experimental bias. Unpublished doctoral dissertation, Ohio State University, 1962.
- GORE, P. M., & ROTTER, J. B. A personality correlate of social action. *Journal of Personality*, 1963, **31**, 58-64.
- GOUGH, H. G. *California Psychological Inventory manual*. Palo Alto: Consulting Psychologists Press, 1964.
- GOUGH, H. G., & HEILBURN, A. B. *The Adjective Check List manual*. Palo Alto: Consulting Psychologists Press, 1965.

- HENDIN, H. Attempted suicide—a psychiatric and statistical study. *Psychiatric Quarterly*, 1950, 24, 39-45.
- HENRY, A. F., & SHORT, J. F., JR. *Suicide and homicide*. Glencoe, Ill.: Free Press, 1951.
- HERSCHE, P. D., & SCHEIBE, K. E. Reliability and validity of internal-external control as a personality dimension. *Journal of Consulting Psychology*, 1967, 31, 609-613.
- HOPKINS, F. Attempted suicide: An investigation. *Journal of Mental Science*, 1937, 83, 71-94.
- JAMES, W. H. Internal versus external control of reinforcement as a basic variable in learning theory. Unpublished doctoral dissertation, Ohio State University, 1957.
- JAMES, W. H., & ROTTER, J. B. Partial and one hundred per cent reinforcement under chance and skill conditions. *Journal of Experimental Psychology*, 1958, 55, 397-403.
- JENKINS, T. N. The accident prone personality. *Personnel*, 1956, 33, 29-32.
- KEEHN, J. D. Factor analysis of reported minor personal mishaps. *Journal of Applied Psychology*, 1959, 43, 311-314.
- KEEHN, J. D. Accident tendency, avoidance learning and perceptual defense. *Australian Journal of Psychology*, 1961, 13, 157-169.
- KUNCE, J. T., & BREWER, B. Neuropsychiatric patients, accident proneness, and interest patterns. *Journal of Psychology*, 1966, 63, 287-290.
- KUNCE, J. T., & WORLEY, B. H. Interest patterns, accidents, and disability. *Journal of Clinical Psychology*, 1966, 22, 105-107.
- LEFCOURT, H. M. Belief in personal control in research and implications. *Journal of Individual Psychology*, 1966, 22, 185-195. (a)
- LEFCOURT, H. M. Internal versus external control of reinforcement. *Psychological Bulletin*, 1966, 65, 206-220. (b)
- LEFCOURT, H. M., & LADWIG, G. W. The effect of reference group upon Negroes' task persistence in a biracial competitive game. *Journal of Personality and Social Psychology*, 1965, 1, 668-671.
- LESHAN, L. L. Dynamics of accident-prone behavior. *Psychiatry*, 1952, 15, 73-80.
- LESHAN, L. L., & BRAME, J. B. A note on techniques in the investigation of accident prone behavior. *Journal of Applied Psychology*, 1953, 37, 79-81.
- LESTER, D. Suicide as an aggressive act. *Journal of Psychology*, 1967, 66, 47-50.
- LITMAN, R. E., & TABACHNICK, N. Fatal one-car accidents. *Psychoanalytic Quarterly*, 1967, 36, 248-259.
- LIVERANT, S., & SCODEL, A. Internal and external control as determinants of decision making under conditions of risk. *Psychological Reports*, 1960, 7, 59-67.
- MACDONALD, J. M. Suicide and homicide by automobile. *American Journal of Psychiatry*, 1964, 121, 366-370.
- MENNINGER, K. A. Purposeful accidents as an expression of self-destructive tendencies. *International Journal of Psychoanalysis*, 1936, 17, 6-16.
- NICKELS, J. B., & TOLEN, E. Use of a basic task in simulating both chance and skill reinforcement. *Proceedings of the 76th Annual Convention of the American Psychological Association*, 1968, 3, 397-398.
- PHARES, E. J. Expectancy changes in skill and chance situations. *Journal of Abnormal and Social Psychology*, 1957, 54, 339-342.
- PHARES, E. J. Perceptual threshold decrements as a function of skill and chance expectancies. *Journal of Psychology*, 1962, 53, 399-407.
- PRESTON, C. E. Accident-proneness in attempted suicide and in automobile accident victims. *Journal of Consulting Psychology*, 1964, 28, 79-82.
- RAPHAEL, T., POWER, S. H., & BERRIDGE, W. L. The question of suicide as a problem in college mental hygiene. *American Journal of Orthopsychiatry*, 1937, 7, 1-14.
- ROBIN, A. A. Review of P. B. Schneider, *La Tentative de Suicide*, 1954. *International Journal of Social Psychiatry*, 1956, 1, 71.
- ROTTER, J. B. *Social learning and clinical psychology*. Englewood Cliffs, N. J.: Prentice-Hall, 1954.
- ROTTER, J. B. Generalized expectancies for internal vs. external control of reinforcement. *Psychological Monographs*, 1966, 80(1, Whole No. 609).
- ROTTER, J. B., LIVERANT, S., & CROWNE, D. P. The growth and extinction of expectancies in chance controlled and skilled tasks. *Journal of Psychology*, 1961, 52, 161-177.
- ROTTER, J. B., & MULRAY, R. C. Internal versus external control of reinforcement and decision time. *Journal of Personality and Social Psychology*, 1965, 2, 598-604.
- SAINSBURY, P. *Suicide in London*. London: Chapman & Hall, 1955.
- SCHNEIDER, J. M. Skill versus chance activity preference and locus of control. *Journal of Consulting and Clinical Psychology*, 1968, 32, 333-337.
- SEEMAN, M. Alienation and social learning in a reformatory. *American Journal of Sociology*, 1963, 69, 270-284.
- SEEMAN, M., & EVANS, J. W. Alienation and learning in a hospital setting. *American Sociological Review*, 1962, 27, 772-783.
- SELZER, M. L., & PAYNE, E. C. Automobile accidents, suicide and unconscious motivation. *American Journal of Psychiatry*, 1962, 119, 237-240.
- SELZER, M. L., PAYNE, E. C., GIFFORD, J., & KELLY, W. Alcoholism, mental illness and the drunk driver. *American Journal of Psychiatry*, 1963, 120, 326-331.
- SHNEIDMAN, E. S. Orientations toward death. In R. W. White (Ed.), *The study of lives*. New York: Prentice-Hall, 1963.
- SHNEIDMAN, E. S. Some reflections on death and suicide. *Folia Psychiatrica et Neurologica Japonica*, 1965, 19, 317-325.
- SEWERS, A. B., & DAVIDOFF, E. Attempted suicide: A survey of 150 patients admitted to two general hospitals. *Journal of Nervous and Mental Disorders*, 1942, 95, 427-441.

- STONEMAN, J. H., & PERTH, W. A. *Halfway and the hereafter*. London: T. F. Christine Ltd., 1935.
- STRICKLAND, B. R. The prediction of social action from a dimension of internal-external control. *Journal of Social Psychology*, 1965, 66, 353-358.
- TABACENICK, N., LITMAN, R. E., OSMAN, M., JONES, W. I., COHN, J., KASPER, A., & MOFFAT, J. Comparative psychiatric study of accidental and suicidal death. *Archives of General Psychiatry*, 1966, 14, 60-68.
- TEICHER, J. D. A study of attempted suicide. *Journal of Nervous and Mental Disorders*, 1947, 105, 283-298.
- TOLOR, A., & REZNIKOFF, M. Relation between insight, repression-sensitization, internal-external control, and death anxiety. *Journal of Abnormal Psychology*, 1967, 72, 426-430.
- VINODA, K. S. Personality characteristics of attempted suicides. *British Journal of Psychiatry*, 1966, 112, 1143-1150.
- WALL, J. H. The psychiatric problem of suicide. *American Journal of Psychiatry*, 1944, 101, 404-406.
- WATSON, D. Relationship between locus of control and anxiety. *Journal of Personality and Social Psychology*, 1967, 6, 91-92.
- WEISS, J. M. A. The gamble with death in attempted suicide. *Psychiatry*, 1957, 20, 17-25.
- WILLIAMS, E. Y. Some observations on the psychological aspects of suicide. *Journal of Abnormal and Social Psychology*, 1936-37, 31, 260-265.
- WINFIELD, D. L., & SPARER, P. J. Preliminary report of the Rosenzweig P-F study in attempted suicides. *Journal of Clinical Psychology*, 1953, 9, 379-381.

(Received August 19, 1968)

Manuscripts Accepted for Publication in the
Journal of Consulting and Clinical Psychology

- Effects of a Model and Instructions on Group Verbal Behaviors: Carol Whalen*: Fairview State Hospital, Costa Mesa, California 92626.
- Effects of Modeling Procedures in Helping Relationships: Kenneth Heller*: Department of Psychiatry, Laboratory of Community Psychiatry, 58 Fenwood Road, Boston, Massachusetts 02115.
- Manipulation of Indications of Disturbed Thinking in Normal Subjects: Leonard S. Levitz* and Leonard P. Ullman: 402 East Green Street, Urbana, Illinois 61801.
- Schizophrenia as a "Put-On": Gerald Rosenbaum*: Department of Psychology, Wayne State University, Detroit, Michigan 48202.
- Schizomimetic Conditions and Schizophrenia: Loren J. Chapman*: Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.
- Staff Ratings of the Relative Importance of the Consequences of Release from or Retention in a Psychiatric Hospital: Theodore W. Lorei*: Veterans Administration Hospital, 50 Irving St., N. W., Washington, D. C. 20422.
- The Hyperkinetic Child: Effect of Incentives on the Speed of Rapid Tapping: Douglas A. Stevens,* Curtis E. Stover, and Joe T. Backus: Phillips County Child Family Guidance Center, 1493 Springdale Road, Helena, Arkansas 72342.
- Hope as a Method of Coping: Ija Dornier*: 1276 Main St., Green Bay, Wisconsin 54302.
- Effects of Grammatical Qualifications on Judgments of the Depth and of the Anxiety Arousal Potential of Interpretive Statements: Edmund S. Howe*: Department of Psychology, University of Missouri, 8001 Natural Bridge Road, Saint Louis, Missouri 63121.
- Effect of Exposure to a Model Receiving Evaluative Feedback on Subsequent Behavior in an Interview: G. Alan Marlatt,* Edward A. Jacobson, Donald L. Johnson, and D. James Morrice: Department of Psychiatry, University of British Columbia, Vancouver 8, Canada.
- Perceived Skill of Professional and Nonprofessional Community Health Workers: Eileen Toban*: 337 East 81 St., New York, New York 10028.
- Professional and Nonprofessional Mental Health Workers' Modes of Persuading Clients to Seek Institutional Services: Eileen Toban*: 337 East 81 St., New York, New York 10028.
- Impact of Audio-Visual Feedback on Group Psychotherapy: Burton N. Danet*: Hamm Memorial Psychiatric Clinic, 238 Hamm Building, Saint Paul, Minnesota 55102.
- Pseudoperceptual Differentiation: Julian Silverman* and Catherine King: Adult Psychiatry Branch, National Institute of Mental Health, Bethesda, Maryland 20014.
- Subprofessional Therapists' Use of Reinforcement Versus Traditional Psychotherapeutic Techniques with Schizophrenics: Lawrence C. Hartlage*: Goodwill Industries, Inc., 1635 W. Michigan St., Indianapolis, Indiana 46222.

* Asterisk indicates author for whom address is supplied.

CHILD-PARENT IDENTIFICATION AND ACADEMIC UNDERACHIEVEMENT

ELNORE GRANLUND AND LOIS KNOWLES¹

Chico State College

The current study extends Shaw and White's investigation of the relationship of child-parent identification and academic underachievement by utilizing a wider range of intelligence scores and a different index of identification. Forty-eight male Ss were classified as either underachievers (UA) or achievers (A) based on the ratio of the prior year's cumulative GPA to their California Test of Mental Maturity IQ score. UA and A masculinity-femininity scores derived from the Bell Adjustment Inventory were significantly different ($p < .01$); no such differences were found between IQ levels or the interaction of IQ \times Achievement. The findings extend Shaw and White's suggestion that male sex-role identification is characteristic of A but not of UA.

A study reported by Shaw and White (1965) indicated a relationship between child-parent identification and academic achievement. Shaw and White found that the high school achiever, but not the underachiever, identifies with the like-sexed parent. In addition, they suggested that an appropriate sex-role identification is characteristic of achievers but not of underachievers.

Since this study was conducted at a single level of intellectual capacity (IQ = 110), the generality of the findings has not been established. The index of identification was based on the correlation between the perceptions of the child and the parent, thus assuming that the identification related to achievement is dependent on identification with the like-sexed parent. The current study extends this investigation to the total IQ range of a small rural high school and the assessment of identification by objective test measures dependent only on the child's interests in and preferences for his sex role. On the basis of the Shaw and White report, it was predicted that (a) underachievers will exhibit significantly lower indexes of appropriate sex-role identification than achievers, and (b) this difference will emerge regardless of level of intellectual capacity.

¹ The authors acknowledge their indebtedness to Richard G. Lathrop of Chico State College for his guidance and counsel. Requests for reprints should be sent to Lois Knowles, 246 North Enright, Willows, California 95988.

METHOD

Subjects

The Ss were selected from the male student population and divided into three equal groups (low, medium, and high IQ) based on IQ scores from the California Test of Mental Maturity. IQ ranges were: low = 74-89; medium = 90-105; high = 106-142. The groups were further subdivided on the basis of ratios of cumulative grade-point average (GPA) for the previous year to IQ score. The lower half of the computed ratios were classified as underachievers; the upper half as achievers. All Ss who were not available for testing or had a GPA from another school district were eliminated from the study. Further random elimination of Ss resulted in the final sample of eight Ss in each of six subgroups (Low, Medium, High IQ \times Underachiever, Achiever)—a total of 48 Ss.

Procedure

The Bell Adjustment Inventory Revised (1962) Student Form (Bell, 1963) was administered to all Ss as a group. The Es indicated that the scores were to be treated in confidence and that none of the results would become a part of the S's school record. The test administration was untimed.

RESULTS

Since Bell (1963) considers a high score on the 30-item masculinity-femininity scale to indicate a good identification with some male during infancy and childhood, scores on this scale were used as an index of sex-role identification. These scores for the six subgroups were analyzed by analysis of variance. The results of this analysis are presented in Table 1. As noted there, achievers and underachievers have significantly different ($p < .01$) masculinity scores. However, no significant differ-

TABLE 1
ANALYSIS OF VARIANCE: MASCULINITY SCORES

Source	df	MS	F
Achievement level	1	136.69	11.24*
IQ (low, medium, high)	2	8.52	.7
Achievement Level \times IQ	2	33.94	2.79
Error	42	12.16	
Total	47		

* $p < .01$.

ences attributable to level of IQ or its interaction with achievement are apparent. Further inspection of the data indicated that males classified as achievers had a higher masculinity score ($\bar{X} = 21.44$) than underachievers ($\bar{X} = 18.08$).

DISCUSSION

The present data indicate a relationship between underachievement and lack of sex-role identification as measured by the Bell Adjustment Inventory's masculinity-femininity scale. When these results are considered in conjunction with the findings of Shaw and White, the hypothesis that identification and

underachievement are related is supported. The current study, however, considerably extends the generality of these conclusions by indicating that this relationship is manifested over the entire range of IQ scores and is not dependent upon a single method of assessing identification. It should be noted, moreover, that the Shaw and White study focuses on parent-child relationships, whereas Bell indicates that his masculinity-femininity scale measures the degree of like-sex identification not necessarily dependent upon parental relationship. Further study is needed to determine whether an appropriate parental identification per se is a prerequisite for academic achievement and whether sex-role identification is a concomitant of appropriate parental identification.

REFERENCES

- BELL, H. M. *Manual—Bell Adjustment Inventory revised (1962) student form*. Palo Alto: Consulting Psychologists Press, 1963.
- SHAW, M. C., & WHITE, D. L. The relationship between child-parent identification and academic underachievement. *Journal of Clinical Psychology*, 1965, 21, 10-13.

(Received April 10, 1968)

WORD ASSOCIATION REPERTOIRES OF SCHIZOPHRENICS AND NORMALS¹

GEORGE D. FULLER² AND SOLIS L. KATES³

University of Massachusetts

The word association repertoires (of three association responses) of 20 poor and 20 good premorbid nonparanoid schizophrenics were not more deviant than those of 20 normals. Over all groups, there was significantly greater commonality in the first association response than in the second response which, in turn, showed greater commonality than the third response. These results suggest that schizophrenic language disorders are not due to overall language deficiency and partially support the disturbance viewpoint. A third viewpoint is proposed, based on a two-stage process involving nondeviant associations with inadequate cognitive controls as differentiating schizophrenics from normals. The hierarchical order for frequency given by the Russell-Jenkins norms obtains for both discrete and extended word association responses.

The present study investigates whether the word association repertoires of poor premorbid and good premorbid nonparanoid schizophrenics are more deviant than those of normal Ss. The association repertoires obtained in this research are made up of three consecutive association responses to each stimulus word.

Several studies using the Kent-Rosanoff word association list with schizophrenics have found that they have associations that are less common than normals (Sommer, Dewar, & Osmond, 1960), more idiosyncratic than normals (Johnson, Weiss, & Zelhart, 1964), and less related to the stimuli than normals (Moran, 1953). On the other hand, the conclusion that schizophrenic patients are not different from normals in their verbal response habits and in their association responses has been reached by other investigators (Dokecki, Polidoro, & Cromwell, 1965; Wynne, 1963). One purpose of this study is to see whether schizophrenic Ss are indeed deviant in their word association responses.

If our results point to greater deviance in word association repertoires for schizophrenic patients, they may also offer some evidence about the two major viewpoints describing

schizophrenic language disorder. One viewpoint suggests that disturbance effects due to competing responses are responsible for schizophrenic deficit (Broen & Storms, 1966) whereas the other viewpoint offers general language deficiency as the responsible factor (Cameron, 1947).

Broen and Storms (1966) infer that, for schizophrenics, competition between the more common and less common responses tends to flatten out the frequency distribution of discrete word associations. Consequently, the initial responses of schizophrenics may show less commonality (are of lower frequency) than those of normals. From this competition-disturbance viewpoint, it is reasonable to assume that common responses which are inhibited as the first association in schizophrenics should tend to emerge as the second or third association in their response hierarchy. If, however, there is an overall language deficiency as inferred by Cameron (1947), schizophrenics should offer association responses that show significantly greater decreases in commonality from the first to the second to the third association than the comparable responses of the normals. Another purpose of this investigation is to determine if there is support for one or the other of the two theoretical systems described above.

In addition, it is predicted that for the combined schizophrenic and normal S groups the hierarchical order of responses in frequency norms for discrete word associations,

¹ Based, in part, on a Master of Science thesis submitted by the senior author in partial fulfillment of the Master of Science degree requirements at the University of Massachusetts.

² Now at Langley Porter Clinic, San Francisco.

³ Requests for reprints should be sent to Solis L. Kates, Department of Psychology, University of Massachusetts, Amherst, Massachusetts 01002.

TABLE 1

MEANS AND STANDARD DEVIATIONS FOR THE MATCHING CRITERIA OF THE SUBJECT GROUPS

Criterion	Group		
	Normals	Good premorbid	Poor premorbid
Age			
<i>M</i>	43.8	44.5	44.0
<i>SD</i>	9.3	7.5	7.2
Education (in grades)			
<i>M</i>	11.1	10.6	11.2
<i>SD</i>	1.8	3.0	2.9
IQ			
<i>M</i>	22.1	21.1	22.4
<i>SD</i>	3.1	2.9	3.4

as described in the Russell-Jenkins (1954) norms, should obtain when word association repertoires of three association responses are secured. Words given as first associations will be represented by a greater frequency on the discrete norms than words given as second associations, and second associations will be represented by a greater frequency than third. It is further predicted that these relationships will hold for both schizophrenic and normal Ss.

METHOD

Subjects

Sixty male Ss in three groups (20 nonparanoid good premorbid schizophrenics, 20 nonparanoid poor premorbid schizophrenics, and 20 normal Ss) were matched on the basis of age, IQ (WAIS Vocabulary and Information), and educational level (Table 1).

Good premorbid were defined as those schizophrenics rated between 0 and 12 points on the Phillips Scale, and poor premorbid as those between 18 and 30 points. As a result of previous studies finding paranoid-nonparanoid to be an independent dimension in schizophrenia (Buss & Lang, 1965), all schizophrenics diagnosed as paranoid or showing significant paranoid features in their case histories were excluded.

Additional criteria for selection of both normal and schizophrenic Ss were: (a) cooperative, attentive normal and schizophrenic males able to understand and conform to the instructions; (b) between the ages of 21 and 60; (c) for schizophrenics, no other complicating pathology (e.g., organicity, mental retardation, alcoholism) and no lobotomy or electroconvulsive therapy within the last 6 months; (d) for normals, presently making an adequate and stable adjustment to the community with no mod-

TABLE 2

MEANS AND STANDARD DEVIATIONS FOR COMMONALITY AND IDIOSYNCRATIC MEASURES

Group	Response			Total
	1	2	3	
Normals				
Commonality				
<i>M</i>	3,982	1,447	911	6,340
<i>SD</i>	1,228	571	466	1,402
Idiosyncratic				
<i>M</i>	3.1	6.6	9.6	19.3
<i>SD</i>	2.2	3.0	2.9	6.3
Good premorbid				
Commonality				
<i>M</i>	3,802	1,263	819	5,884
<i>SD</i>	1,196	615	563	1,575
Idiosyncratic				
<i>M</i>	2.8	8.8	10.2	21.8
<i>SD</i>	2.7	3.2	4.1	8.5
Poor premorbid				
Commonality				
<i>M</i>	3,516	1,146	682	5,344
<i>SD</i>	1,202	601	459	1,638
Idiosyncratic				
<i>M</i>	3.4	7.8	10.8	22.0
<i>SD</i>	2.3	3.7	3.7	8.4

erate or disabling emotional disturbance or alcoholism.

Test Materials

Twenty-four words were selected from the Kent-Rosanoff word association list which were likely to yield superordinate association responses (e.g., "fruit" as the superordinate or class response for the stimulus word "apple"). The stimulus words were presented to Ss singly on 3 × 5 inch cards.

Procedure

The Ss were tested individually and the evaluation of each S's repertoire was based on the scoring of three consecutive responses to each stimulus word.

The groups were compared on (a) commonality scores where each association given by S was assigned a score equal to its response frequency as given in the Russell-Jenkins (1954) norms and (b) the number of idiosyncratic responses (associations given by Ss having a frequency of zero in the above norms).

RESULTS

Analyses of variance procedures for both the commonality of response data and the number of idiosyncratic associations showed that the poor premorbid, good premorbid, and

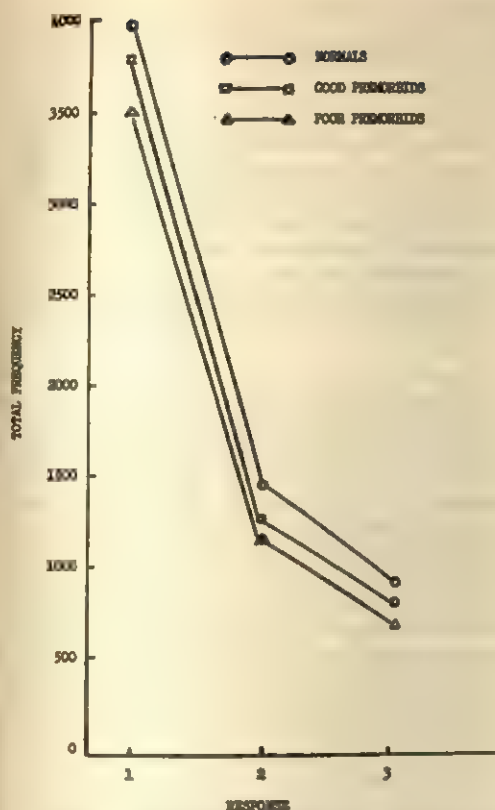


FIG. 1. Performance of groups on commonality measure for first three responses.

normal groups did not differ significantly from each other.

The three responses differed significantly from each other on the commonality measure ($F = 238$, $df = 2/108$, $p < .001$) and on the idiosyncratic measure ($F = 180$, $df = 2/108$, $p < .001$) (Table 2). For each dependent measure, the first association response differed significantly from the second which, in turn, differed significantly from the third response. The three diagnostic groups show closely parallel response curves for the commonality and idiosyncratic measures which are quite clearly seen in Figures 1 and 2. No significant interactions between responses and diagnostic groups were noted for commonality and idiosyncratic scores.

DISCUSSION

The results show that nonparanoid schizophrenics, both good and poor premorbid, possess word association response repertoires

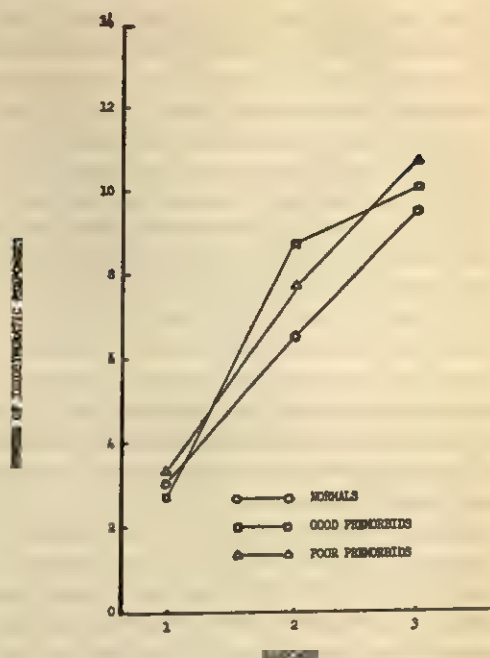


FIG. 2. Performance of groups on idiosyncratic measure for first three responses.

that are very similar to those of normal Ss. First, no significant differences were noted among the groups on the measures of commonality and idiosyncratic responses given to the 24 selected Kent-Rosanoff words. Further, the response curves for commonality and idiosyncratic responses are very similar for poor and good schizophrenic and normal Ss when their three association responses are compared separately. These findings question the viewpoint that schizophrenics have an overall language deficiency. On the other hand, our findings support the disturbance viewpoint, in part, and affirm the inference that the first, second, and third word-association responses in schizophrenics do not show a progressive decrease in commonality greater than the corresponding responses of normals.

Possibly a different account of schizophrenic disorder may be proposed based upon the highly similar two-stage approach independently suggested by Cohen and Camhi (1967) and also by Schwartz and Rouse (1961). Cohen and Camhi (1967) reject the conclusion that schizophrenics have deviant word association repertoires from which they sample

responses, and assert that they make inappropriate selections from these nondeviant repertoires, which normals are better able to suppress. Schwartz and Rouse (1961) divide the association process into associative linkages and cognitive controls, holding that in clinical patients there is an abnormal modification of the cognitive controls. The associative linkages in schizophrenic patients are probably similar to those of the normals, but it is our suggestion that there is an impairment of the cognitive controls in schizophrenics permitting associations to emerge in contexts where they are inappropriate. These associations, derived from nondeviant association repertoires, are consequently deviant when overtly expressed.

The finding that the initial association response of both schizophrenic and normal Ss was more common and less idiosyncratic than the second response which, in turn, was more common and less idiosyncratic than the third response for the diagnostic groups supports the Russell-Jenkins (1954) norms for commonality and idiosyncrasy. It further indicates that the hierarchical order of responses in the frequency norms for discrete word associations obtains for extended word associations involving three responses.

REFERENCES

- BROEN, W. E., JR., & STORMS, L. H. Lawful disorganization of the process underlying the schizophrenic syndrome. *Psychological Review*, 1966, 73, 265-279.
- BUSS, A. H., & LANG, P. J. Psychological deficit in schizophrenia: I. Affect, reinforcement, and concept attainment. *Journal of Abnormal Psychology*, 1965, 70, 2-24.
- CAMERON, M. *The psychology of behavior disorders*. Boston: Houghton Mifflin, 1947.
- COHEN, B. D., & CAMHI, G. Schizophrenic performance in a word communication task. *Journal of Abnormal Psychology*, 1967, 72, 240-246.
- DOKECKI, P. R., POLIDORO, L. G., & CROMWELL, R. L. Commonality and stability of word association responses in good and poor premorbid schizophrenics. *Journal of Abnormal Psychology*, 1965, 70, 312-316.
- JOHNSON, R. C., WEISS, R. L., & ZELHART, P. F. Similarities and differences between normal and psychotic subjects in response to verbal stimuli. *Journal of Abnormal and Social Psychology*, 1964, 68, 221-226.
- MORAN, L. J. Vocabulary knowledge and usage among normal and schizophrenic subjects. *Psychological Monographs*, 1953, 67(20, Whole No. 370).
- RUSSELL, W., & JENKINS, J. J. *The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test*. University of Minnesota studies on the role of language in behavior. Technical Report No. 11, August 1954. Office of Naval Research.
- SCHWARTZ, F., & ROUSE, R. O. The activation and recovery of associations. *Psychological Issues*, 1961, 3(Monogr. 9).
- SOMMER, R., DEWAR, R., & OSMOND, H. Is there a schizophrenic language? *Archives of General Psychiatry*, 1960, 3, 665-673.
- WYNNE, R. D. Can schizophrenics give the associations that "most people" do? Paper read at the meeting of the Eastern Psychological Association, New York, April 1963.

(Received June 27, 1968)

INTELLIGENCE OF VOLUNTEERS AS RESEARCH SUBJECTS¹

JAY M. CUDRIN²

Connecticut Valley Hospital, Middletown

The personal characteristics of people who volunteer for research is becoming increasingly important to behavioral scientists. This study compared the intelligence of prison inmates who volunteered for medical experiments with that of inmates who did not volunteer. The results showed that groups of volunteers for two individual experiments demonstrated significantly higher scores on the Revised Beta Examination than did nonvolunteers. The implications of these findings to behavioral research were examined.

Recent inquiries into invasion of privacy through the use of psychological tests (Brayfield, 1965) and research in the social sciences (Brayfield, 1967) have made it increasingly important for psychologists to carry out their research with voluntary Ss. This emphasis on the use of volunteers underscores the necessity for examining the personal characteristics of such Ss. A growing body of literature has already begun to scrutinize this issue. In their discussion of the Kinsey study, for example, Maslow and Sakoda (1952) showed that volunteers tended to have higher self-esteem than nonvolunteers. Riggs and Kaess (1955) learned that when students in a psychology class were asked to volunteer to take a psychological test, those who came forward showed more introverted thinking and moodiness than nonvolunteers. The volunteers also demonstrated greater cyclical trends in emotionality, higher intropunitiveness, and lower extrapunitiveness. This entire issue has also been studied and discussed by Bair and Gallagher (1960), Locke (1954), and Newman (1956).

In view of the number of volunteer characteristics already studied, it is noteworthy that no one has examined the relative intelligence level of volunteers as a group. Investigators generally recognize the importance of intelligence in behavioral research, and usually try to equate the intellectual levels of their experimental and control groups. It may be that people volunteering for a study have a

different intelligence level than that of the population from which they are drawn. This would mean that, in spite of equating experimental and control groups, decisions would still have to be made about the appropriateness of generalizing results to the population at large.

The purpose of this study was to learn whether people who volunteer for research do differ intellectually from the rest of the population. Since psychological research implies to the ordinary individual some kind of mental activity, the chances of volunteers being brighter than nonvolunteers is increased. That is, people who realistically see themselves as less intelligent might be reluctant to embarrass themselves by volunteering for a "psychological experiment." This study, therefore, looked at volunteers who came forward for a project which had no implication whatever of examining mental behavior. The volunteers had offered their services for a medical project. We were thus able to learn something about whether the mere idea of volunteering (for some kind of research) appealed more to people of a certain intellectual stratum.

METHOD

The study was performed at a federal prison in the southeastern United States where public health physicians were engaged in two medical projects: one on malaria and the other on gonorrhea. Each required that volunteers artificially contract the disease, and then stay in the prison infirmary until cured.

Volunteers for the projects were obtained from the prison's inmate population. All inmates were told about these projects and the need for volunteers during their first week at the prison. They were told they could volunteer any time during their sentence. The only conditions they had to meet were to pass a routine physical examination and to have shown

¹ This paper was presented at the meeting of the Eastern Psychological Association, Philadelphia, April 1969.

² Requests for reprints should be sent to Jay M. Cudrin, Department of Psychology, Connecticut Valley Hospital, Middletown, Connecticut 06457.

little evidence of severe disciplinary infractions at the prison. The latter was an attempt to avoid behavioral problems in the hospital. Few men were ruled out on the basis of it.

When each man finished the project in which he was participating, he received a small stipend. A letter was also placed in his record indicating that he had taken part in the experiment. He was informed, though, that the letter constituted neither a pardon nor an assurance of parole.

The data for this study consisted of inmates' scores on the Revised Beta Examination. Forty men who had volunteered for the malaria study had taken this test as part of the prison's routine admission procedure. An equal number from the gonorrhea study was obtained by eliminating every fourth score from an alphabetical list of volunteers' names (totaling 49) until 40 scores were left. Finally, 40 nonvolunteers' scores were gotten by taking every fourth score from an alphabetical list of the prison population. There was no overlap in any group.

RESULTS

The mean Beta scores and standard deviations of the three groups were: for malaria, $M = 106.65$, $SD = 12.84$; for gonorrhea, $M = 108.90$, $SD = 9.11$; for nonvolunteers $M = 100.32$, $SD = 11.49$.

The significance of differences between means of each of the three groups was determined by t tests. The results of the t tests are shown in Table 1 and indicate that volunteers manifest significantly higher scores than do nonvolunteers.

DISCUSSION

The results of this study show that inmate volunteers had significantly higher IQ scores than the population from which they came. There is the issue of whether the IQ differences examined here are clinically as well as statistically significant. It would be somewhat difficult to make a case for functional differences if only one experimental group were tested. This study, however, compared IQ means from each of two independent groups of volunteers with the IQ mean of nonvolunteers. Each experimental group was involved in a separate research project. The fact that identical findings were obtained each time shows that, although the IQ differences were not large, volunteers consistently came from a significantly different population. Thus, it is the consistency of the findings that adds

TABLE 1

SIGNIFICANCE OF DIFFERENCES BETWEEN REVISED BETA MEANS FOR EACH OF THREE GROUPS

Groups	t	df	p
Malaria vs. nonvolunteers	2.36	78	.05
Gonorrhea vs. nonvolunteers	3.60	78	.01
Malaria vs. gonorrhea	.94	78	<i>ns</i>

weight to the conclusion that the groups were functionally different.

The author can suggest two possible reasons for why volunteers came from a more intelligent population. One is that although inmates were given a certain amount of assurance that neither experiment would jeopardize their long range health, only the more intelligent men may have really accepted the idea they would not be harmed. An even more plausible hypothesis grew out of a conversation the author had with one of the volunteers. The man said he had volunteered in order to give himself something interesting to do. Experiences like these experiments, he said, helped break up the dullness of his sentence. If other volunteers shared this point of view, a desire for stimulation and unique experience was important in motivating men to participate. Of course, prison life is the sort of experience that is apt to increase the need for psychological stimulation. The implicit correlation found here between intelligence and desire to participate in uniquely different experiences is not different from the findings of other researchers (Keniston, 1968; Terman & Oden, 1947).

Other variables related to intelligence may also have been influential here. Education and socioeconomic level are just two examples. The volunteers may have been men of higher formal education, and schooling could have predisposed them to taking part in research efforts. In addition, the brighter men may have come from a generally higher socioeconomic level. This social background may have left them, in spite of their criminal behavior, with more of the social awareness that leads to volunteering for research projects. Unfortunately, the only data the author had access to were the intelligence test scores.

Thus, there are several hypotheses which still could be profitably investigated.

For the moment though the results do show a clear and consistently positive relationship between the volunteering of prison inmates to be research Ss and level of intelligence. It would be imprudent to automatically generalize these findings to all groups of volunteer Ss. The conditions of the situation for which the men here offered themselves, however, do give psychologists reason for thought: if simply volunteering for research which implies no mental activity whatever draws on a somewhat brighter than average group of people, we have reason to be concerned over the intellectual representativeness of volunteers for psychological studies.

REFERENCES

- Bair, J. T., & Gallagher, T. J. Volunteering for extra-hazardous duty. *Journal of Applied Psychology*, 1960, 44, 329-331.
- Brayfield, A. H. (Ed.) Testing and public policy. *American Psychologist*, 1965, 20, 857-992.
- Brayfield, A. H. (Ed.) Congress and social science. *American Psychologist*, 1967, 22, 877-1035.
- Keniston, K. Psychological aspects of use and abuse of marijuana. Paper presented at the meeting of the Connecticut Valley Section of the American Chemical Society, West Hartford, November 1968.
- Locke, H. J. Are volunteer interviewees representative? *Social Problems*, 1954, 1, 143-146.
- Maslow, A. H., & Sakoda, J. M. Volunteer error in the Kinsey study. *Journal of Abnormal and Social Psychology*, 1952, 47, 259-262.
- Newman, M. *Personality differences between volunteers and non-volunteers for psychological investigations; self-actualisation of volunteers and non-volunteers for researches in personality and perception*. (Doctoral dissertation, New York University) Ann Arbor, Mich.: University Microfilms, 1956. No. 19,999.
- Riggs, M. M., & Kaess, W. Personality differences between volunteers and non-volunteers. *Journal of Psychology*, 1955, 40, 229-245.
- Terman, L. M., & Oden, M. H. *The gifted child grows up*. Stanford: Stanford University Press, 1947.

(Received August 1, 1968)

(Continued from page 494)

- Performance of Normal and Organic Brain-Damaged Subjects on the Canter Background Procedure (BIP) Test as a Function of Drive: Sergio Yulis*: Psychology Department, Norristown State Hospital, Norristown, Pennsylvania 19401.
- Reactions to the Names "Counseling Center" and "Psychological Center": Nick A. Sieveking* and James E. Chappell: Psychological and Counseling Center, Vanderbilt University, Nashville, Tennessee 37203.
- Level of Aspiration, Risk-Taking Behavior, and Projective Test Performance: A Search for Coherence: Herbert M. Lefcourt* and Richard A. Steffy: Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada.
- Relationships between Global and Specified Measures of Novelty Seeking: Pamela H. Pearson*: Department of Psychology, University of Chicago, Chicago, Illinois 60637.
- Psychodiagnostic Training in the Academic Setting: Past and Present: Kenneth Shemberg* and Stuart Keeley: Department of Psychology, Bowling Green State University, Bowling Green, Ohio 43402.
- Effect of Intelligence, Anxiety, and Diagnosis on Arithmetic and Digit Span Performance on the WAIS: Charlotte J. Callens and Malcolm L. Meltzer*: Department of Psychology, George Washington University, Washington, D. C. 20006.
- Sixteen Personality Factor Questionnaire Scores of Homosexual Men: Ray B. Evans*: Department of Psychiatry, Loma Linda University, Loma Linda, California 92354. ■ ■ ■
- Invitation of Punishment by Excessive Drinkers in Treatment: Stanley V. Butts* and Franklin C. Shontz: Department of Nursing Education, University of Kansas Medical Center, Rainbow Boulevard at 39th St., Kansas City, Kansas 66103.
- Dimensions of Internal Versus External Control: Herbert L. Mirels*: Department of Psychology, Ohio State University, 1945 North High Street, Columbus, Ohio 43210.
- Anxiety, Structure, or SD?: Allen L. Edwards*, John D. Cone, Jr., and Robert D. Abbott: Department of Psychology, University of Washington, Seattle, Washington 98105.
- Therapists' Responses as a Function of Level of Therapist Experience and Attitude of the Patient: Judith Williams Beery*: Counseling and Testing Center, Stanford University, 306 Student Services Center, Stanford, California 94305.
- Field Independence and Educational-Vocational Interests: Jack Arbuthnot* and Leopold Gruenfeld: New York State School of Industrial and Labor Relations, Cornell University, Ithaca, New York 14850.

* Asterisk indicates author for whom address is supplied.

HUMOR AND SUICIDE:

FAVORITE JOKES OF SUICIDAL PATIENTS¹

DON SPIEGEL,² PATRICIA KEITH-SPIEGEL, JOEL ABRAHAMS

AND LIONEL KRANITZ

Brentwood Hospital, Veterans Administration Center, Los Angeles

During an interview 20 patients who had attempted suicide and 20 matched nonsuicidal patients were asked to relate their favorite joke. Jokes were told by 12 attempters and 9 controls. Two raters, blind to the purpose and nature of the study, rated each joke for direction of punishment expressed in the themes. Based on the combined chi-square probabilities associated with ratings by the two judges, suicidal Ss told significantly more jokes with a self-punishing theme than did nonsuicidal controls.

The technique of using jokes in an attempt to provide clinical insights into motivational states on the basis of the assumption that a person unconsciously identifies with the central figure in his favorite joke has been described in the literature. Brill (1940) felt that the favorite joke was used to express hostile and other tendencies for which direct expression is restricted or entirely shut off, and that the content of the joke fits the patient's struggle and furnishes him a certain amount of pleasure—whether a certain freedom from anxiety or a discharge of tension—by making light of his difficulties.

More recently, Zwerling (1955) noted that the very themes that were most provocative of the patient's anxieties were precisely the subjects of jokes, as though to laugh at something is to deny that it arouses anxiety. Here the joke may function, as Freud (1928) wrote, as the loftiest of the defense mechanisms available to man.

Research studies in which the favorite joke technique is used to assess group differences appear to be nonexistent. The majority of humor studies have focused on the responses of individuals or groups to the visual or auditory presentation of humor stimuli.

In the present study the assumption was made that the self-destructive tendencies of suicidal individuals tend to manifest them-

selves in various facets of their behavior, including humor productions. It was further assumed that people tend to identify with the central figure of their favorite jokes. Thus, it was predicted that the thematic content of favorite jokes of suicide attempters would more frequently involve harm to the central figure than would the favorite jokes of nonsuicidal controls.

METHOD

The Ss were 20 patients consecutively admitted to a neuropsychiatric hospital over a 2-month period because of suicide attempts and 20 patients from the same facility who were matched individually for diagnosis, age, and sex, but who had no history of suicide attempts, threats, or ideation. Each group contained 17 males and 3 females between 29 and 56 years of age. In each group were nine schizophrenic, nine depressive, and two alcoholic patients. Lengths of hospitalization and acuteness-chronicity in the two groups were similar.

All 40 Ss were given a standard interview and then asked to tell their favorite joke. Their productions were recorded verbatim by the interviewer. Jokes were told by 12 attempters and 9 controls. These were typed on separate sheets of paper, assigned code numbers, and given to two psychologists to rate independently for direction of aggression or harm expressed in the jokes, using a system derived from the work of Rosenzweig (1945). Raters were blind to the purpose of the study. If the hero (or general theme) of the joke was outwardly aggressive, the joke was classified as *extropunitive*. Actual jokes given by patients are used here to exemplify the categories into which jokes were sorted. (Example: "What are the things between elephants' toes? Slow natives.") If the hero or main character(s) suffered harm, pain, or discomfort, the joke was classified as *intropunitive*. (Example: "What did the firefly say as he backed into the fire? Delighted, no end.")

¹ Lisa Gray Shellberg and Barbara S. Eaton served as raters for the study.

² Requests for reprints should be sent to Don Spiegel, Brentwood Hospital (183 B), Veterans Administration Center, Wilshire and Sawtelle Boulevards, Los Angeles, California, 90073.

If there was no aggressive or punitive theme, and if no harm or discomfort was suffered by any character in the joke, it was classified as *impunitive*. (Example: "Why does a fireman cross the street? To get to the other side.")

RESULTS

The following frequencies of jokes assigned to categories for each patient group are listed in the order, intropunitive, impunitive, and extropunitive, with frequencies for controls in parentheses: Judge A: 7(0), 4(6), 1(3), and Judge B: 5(1), 2(5), 5(3). Frequencies for the impunitive and extropunitive categories were then combined into an "other" category, and for each independent rater, the chi-square method was used to test the hypothesis that more suicide attempters than controls told jokes with intropunitive content. The obtained chi-square values for each rater were: Judge A: 5.47 ($p < .01$), Judge B: 1.09 ($p < .15$). Using the method described by Gordon, Loveland, and Cureton (1952), a combined chi-square value of 13.03 was obtained which was significant at $p < .02$.

DISCUSSION

The fact that statistical significance was obtained with such small groups of joke tellers is impressive in itself. However, since only slightly over half the patients produced a joke on request, and since it is not always

possible to tell whether a joke produced has some personal significance for the teller (or is merely the most recent joke he has heard or the only one he can remember), this technique can at best serve as an adjunct to more reliable diagnostic procedures. Clinically, it would appear wise to attempt to obtain as much information as possible about the meaning of the joke to the patient, why it is his favorite joke, when and how he learned it, etc. Thus, an individual's humor productions, like any other potentially projective material, may be considered significant only to the extent that it corroborates other findings in an assessment procedure.

REFERENCES

- BRILL, A. The mechanism of wit and humor in normal and psychopathic states. *Psychiatric Quarterly*, 1940, 14, 731-749.
- FREUD, S. On humor. *International Journal of Psychoanalysis*, 1928, 9, 1-6.
- GORDON, M. H., LOVELAND, E. H., & CURETON, E. E. An extended table of chi square for two degrees of freedom, for use in combining probabilities from independent samples. *Psychometrika*, 1952, 17, 311-326.
- ROSENZWEIG, S. The picture-association method and its application in a study of reaction to frustration. *Journal of Personality*, 1945, 14, 3-23.
- ZWERLING, I. The favorite joke in diagnostic and therapeutic interviewing. *Psychoanalytic Quarterly*, 1955, 24, 104-114.

(Received August 1, 1968)

BRIEF REPORTS

MOTIVATIONAL PROPERTIES OF PERSONALITY CONFLICT¹

RONALD E. SMITH

University of Washington

A number of S-R theorists have suggested that conflict induces a drive state, assumed to have energizing properties, over and above the original ones motivating conflict behavior in a given situation. Brown and Farber (1968) have reviewed a number of supportive studies indicating that conflict may result in an energization of ongoing behavior and that escape from conflict is rewarding. A method for assessing personality conflict in a number of need areas has been proposed by Worell and Worell (1965). They assume that conflict is indicated when an individual's stated behavior choices differ from his stated behavior preferences, for he is indicating that competing dispositions are replacing supposedly preferred alternatives. The relatively nonconflicted individual, on the other hand, tends to use the behaviors that he states he prefers. The Conflict Scale (Hays & Worell, 1967), designed to measure this preference-behavior differential, is composed of modified scales from the Edwards Personal Preference Schedule. The present study employed this scale to assess the effects of conflict in the need achievement area on the learning of noncompetitive and competitive paired associates under achievement-oriented conditions. The Ss were male college students divided into high- and low-conflict groups, each composed of 14 Ss, on the basis of their Conflict Scale scores. All of the Ss were above the median in achievement preference. The Ss were individually tested on noncompetitive and competitive lists of paired

associates. They were told that the tasks measured important concept formation skills and that they would be given information regarding their level of performance. On the basis of drive theory, it was predicted that if conflict increases drive level, the high-conflict group would outperform the low group on the noncompetitive list while performing relatively more poorly on the competitive list. The mean number of correct anticipations for the two conflict groups on the nine-item noncompetitive list, which was presented five times, were: high, 34.29; low, 28.14. On the eight-item competitive list, presented six times, the means were: high, 21.29; low, 22.64. Analysis of variance yielded a significant Conflict \times List interaction ($F = 6.16$, $df = 1/26$, $p < .02$), confirming the hypothesis. The role of conflict in the performance differences of the two conflict groups is further supported by the fact that the groups did not differ significantly in TAT n Achievement or Test Anxiety, the two motivational variables most likely to affect performance in such a situation. The mean TAT n Achievement scores for the conflict groups were: high, 4.57; low, 7.64 ($t = 1.19$, $df = 26$). The mean Test Anxiety scores for the two conflict groups were: high, 69.43; low, 72.71 ($t = .23$, $df = 26$). The results are interpreted as supporting the notion that conflict, defined as the simultaneous arousal of incompatible response tendencies, has motivational (drive) properties associated with it.

REFERENCE

- ¹An extended report of this study may be obtained without charge from Ronald E. Smith, Department of Psychology, University of Washington, Seattle, Washington 98105, or for a fee from the National Auxiliary Publications Service. Order Document No. 00513 from the National Auxiliary Publications Service of the American Society for Information Science, c/o CCM Information Sciences, Inc., 909 Third Avenue, New York, New York 10022. Remit in advance \$3.00 for photocopies or \$1.00 for microfiche and make checks payable to: Research and Microfilm Publications, Inc. Requests for reprints should be sent to the author at the above address.
- BROWN, J., & FARBER, I. E. Secondary motivational systems. In P. R. Farnsworth (Ed.), *Annual review of psychology*. Palo Alto: Annual Reviews, 1968.
- HAYS, L., & WORELL, L. Anxiety drive, anxiety habit, and achievement: A reformulation in terms of optimal motivation. *Journal of Personality*, 1967, 35, 145-161.
- WORELL, J., & WORELL, L. Personality conflict, originality of response, and recall. *Journal of Consulting Psychology*, 1965, 29, 55-62.

(Received November 21, 1968)

PERCEIVED MATERNAL NURTURANCE AND CONTROL OF PROCESS SCHIZOPHRENICS, REACTIVE SCHIZOPHRENICS, AND NORMALS¹

PETER A. MAGARO

AND

BARBARA A. HANSON

University of Maine

Northern Illinois University

The present study attempts to compare the amount of maternal nurturance and control as measured by Heilbrun (1964) in three S groups—process schizophrenics, reactive schizophrenics, and normals. It has been reported that mothers of process schizophrenics tend to be more domineering, restricting, rejecting, and demanding of their sons than the mothers of reactive schizophrenics or normals (Rodnick & Garnezy, 1957). Considering these findings and the usual theorizing of differential child-rearing histories for normals and schizophrenics, it was predicted that the process schizophrenics would rate their mothers as exhibiting less nurturance and more control than the reactive schizophrenic or normals.

The Ss included 198 male undergraduate students randomly selected from introductory psychology classes, 33 male process schizophrenics, and 33 male reactive schizophrenics. The process-reactive breakdown was made on the basis of the Phillips Scale of Premorbid Adjustment (Phillips, 1953). The tests administered were the Parent Attitude Research Instrument (Schaefer & Bell, 1958) measuring control and the Parent-Child Interaction Rating Scale (Heilbrun, 1964) measuring nurturance.

As predicted, there was a significant difference between the three groups on the nurturance fac-

tor ($F = 4.71$, $df = 2/261$, $p < .01$). The process group exhibited the lowest nurturance scores, differing significantly from the reactive and normal groups who did not differ from one another. A difference between the normals and schizophrenics was found on the control scale with both schizophrenic groups exhibiting higher control than normals. However, the design of the Bell Scale, involving a multiple selection of responses, was felt to be subject to a response set with schizophrenic Ss. Therefore, the control scale was not seen as being a clear indicator of maternal control.

It is felt that these results support the hypothesis of nurturance differences being present between the mothers of process schizophrenics and mothers of reactive schizophrenics. Process schizophrenics report the child-rearing period as devoid of affection, while reactive schizophrenics, who had a better premorbid adjustment, rated their mothers similar to normals, that is, higher in demonstrated affection. The results also tend to support the process-reactive distinction in regard to differences in prior family interaction.

REFERENCES

- HEILBRUN, A. B. Parent model attributes, nurturant reinforcement and consistency of behavior in adolescents. *Child Development*, 1964, 35, 151-167.
- PHILLIPS, L. Case history data and prognosis in schizophrenia. *Journal of Nervous and Mental Disease*, 1953, 117, 515-525.
- RODNICK, E. H., & GARMEZY, N. An experimental approach to the study of motivation in schizophrenia. *Nebraska Symposium on Motivation*, 1957, 5, 109-184.
- SCHAEFER, E. S., & BELL, R. A. Development of a parent attitude research instrument. *Child Development*, 1958, 29, 339-361.

(Received November 25, 1968)

¹ An extended report of this study may be obtained without charge from Peter A. Magaro, Psychology Department, University of Maine, Orono, Maine 04473, or for a fee from the National Auxiliary Publications Service. Order Document No. 00514 from the National Auxiliary Publications Service of the American Society for Information Science, c/o CCM Information Sciences, Inc., 909 Third Avenue, New York, New York 10022. Remit in advance \$3.00 for photocopies or \$1.00 for microfiche and make checks payable to: Research and Microfilm Publications, Inc. Requests for reprints should be sent to Peter A. Magaro at the above address.

PERFORMANCE CHARACTERISTICS OF SCHOOL CHILDREN ON THE BIP BENDER TEST¹

JERRY ADAMS

AND

ARTHUR CANTER

William S. Hall Psychiatric Institute

University of Iowa Psychiatric Hospital

The Background Interference Procedure (BIP) as applied to the Bender Test for the detection of organic brain disorder has been developed for use with adults. However, there appears to be an insistent need for such a technique that may be used with children. The purpose of the present study was to determine what the "normal" characteristics of children's performance would be on the BIP Bender.

A total of 272 presumably normal children, sampled from rural and small city schools, ranging in age from 6 to 14 years were administered the BIP Bender Test and a Coding test based on the one in the WISC battery. Mental ability test scores on the children were also available. The BIP Bender tests were scored by the method proposed by Canter (1968) which yields four scores: error scores for both standard Bender and BIP test administrations, and two BIP decrement indexes, a D score, and "number positive" (NP) representing the difference in total errors between the two test administrations and the number of designs worse on BIP, respectively.

It was found that the scores on the Bender, BIP, and Coding tests changed with age. The

Bender and BIP error scores decreased fairly regularly with age whereas the BIP D scores decreased sharply around age 8 and tended to vary more narrowly after age 10. The correlation between BIP D score and age was $-.25$, a significant ($p < .01$) but low-order correlation. Extremes of the intelligence range in the sample were compared on BIP D scores and usual Bender error scores. Whereas the latter showed significant differences between intelligence levels, the BIP D scores did not. Also no significant relationship was found between Coding scores and the BIP D scores. No sex differences were observed on the various BIP Bender scores.

It would appear that the ability to cope with the BIP effect, as measured by the D score and NP scores, is fairly well established by the age of 13 and is relatively unaffected by differences in intelligence. Below this age, greater allowances need to be made for what appears to be a normal decrement in performance associated with the BIP task. This is especially true for children under 8 years of age. Thus one should not apply Canter's criteria for brain-damage performance on the BIP Bender to children below the age of 13. The way the procedure is used with adults as well as some of the types of error deviations measured on the Bender probably need to be modified to make the technique more useful for the evaluation of children. This is a matter for further investigation.

REFERENCE

- CANTER, A. BIP Bender test for the detection of organic brain disorder. *Journal of Consulting and Clinical Psychology*, 1968, 32, 522-526.

(Received December 3, 1968)

¹An extended report of this study may be obtained without charge from Arthur Canter, University of Iowa Psychiatric Hospital, 500 Newton Road, Iowa City, Iowa 52240, or for a fee from the National Auxiliary Publication Service. Order Document No. 00515 from the National Auxiliary Publications Service of the American Society for Information Science, c/o CCM Information Sciences, Inc., 909 Third Avenue, New York, New York 10022. Remit in advance \$3.00 for photocopies or \$1.00 for microfiche and make checks payable to: Research and Microfilm Publications, Inc. Requests for reprints should be sent to Arthur Canter at the above address.

JOURNAL OF CONSULTING and CLINICAL PSYCHOLOGY

October 1969

Vol. 33, No. 5

Contents

ARTICLES

Effects of a Model and Instructions on Group Verbal Behaviors: <i>Carol Whalen</i>	509
Effects of Modeling Procedures in Helping Relationships: <i>Kenneth Heller</i>	522
Conditioning and Transfer of Affective Self-References in a Role-Played Counseling Interview: <i>Robert J. Hoffnung</i>	527
Adjustment, Mental Health Opinions, and Proficiency of Child Care Personnel: <i>John W. Giebink and Donald O. Stover</i>	532
Therapist Directiveness, Client Dominance, and Therapy Resistance: <i>Robert W. Hagebak and George V. C. Parker</i>	536
Temporal Correlates of the Achievement Value and Manifest Anxiety: <i>Thomas J. Cottle</i>	541
Different Psychological Effects of Lateralized Brain Damage: <i>Oscar A. Parsons, Arthur Vega, Jr., and Julian Burn</i>	551
Relationship of Academic Underachievement to Incidental Learning: <i>Michele S. Mondani and Thomas A. Tutko</i>	558
Self-Care Unit: Some Lessons in Institutional Power: <i>Alan P. Towbin</i>	561
Sensation Seeking, Impulsivity, and Psychopathic Personality: <i>R. Blackburn</i>	571
Personality Characteristics of Parents of Neurotic, Aggressive, and Normal Preadolescent Boys: <i>Lorna M. Anderson</i>	575
Locus of Control in Psychiatric Patients: <i>Martin Harrow and Anthony Ferrante</i>	582
Study of Psychological Content in the Communications of Suicidal Individuals: <i>Allen R. Darbonne</i>	590
Predicting the Success of Peace Corps Community Development Workers: <i>Charles Dicken</i>	597
Comparison of the Binet and WPPSI with Lower-Class Five-Year-Olds: <i>Joan Fagan, Elizabeth Broughton, Mildred Allen, Betty Clark, and Patricia Emerson</i>	607
Participation in Psychological Research: Relation to Birth Order and Demographic Factors: <i>Edward H. Fischer and David Winer</i>	610
Direction of Affective Influence in Psychotherapy: <i>Kenneth I. Howard, Merton S. Krause, and David E. Orlinsky</i>	614
Prediction Enhancement through the Use of Moderator Variables: <i>James E. Cowden</i> ..	621
Sex Differences and Variations in Body Attitudes: <i>Richard M. Kurtz</i>	625

BRIEF REPORTS

Effect of Intelligence, Anxiety, and Diagnosis on Arithmetic and Digit Span Performance on the WAIS: <i>Charlotte J. Callens and Malcolm L. Meltzer</i>	630
Field Independence and Educational-Vocational Interests: <i>Jack Arbuthnot and Leopold Gruenfeld</i>	631
Impact of Audio-Visual Feedback on Group Psychotherapy: <i>Burton N. Danet</i>	632

LIST OF MANUSCRIPTS ACCEPTED.....	581, 589, 620
-----------------------------------	---------------

Brief Reports

The *Journal of Consulting and Clinical Psychology* will accept Brief Reports of research studies in clinical psychology for early publication without expense to the author. The procedure is intended to permit the publication of soundly designed studies of specialized interest or limited importance which cannot now be accepted because of lack of space. Several pages in each issue will be devoted to Brief Reports, published in the order of their receipt without respect to the dates of receipt of the regular articles. Most Brief Reports appear in the first or second issue to go to press following their final acceptance.

An author who wishes to submit a Brief Report:

1. Sends the Brief Report, limited to one printed page and prepared according to the specifications given below.

2. Also sends to the Editor a full report of the research study, in sufficient detail to give a clear account of its background, procedure, results, and conclusions, which will be filed with the National Auxiliary Publications Service (NAPS) to insure indefinite availability.

3. Prepares at least 100 mimeographed copies of the full report, which the author will send without charge to all who request it as long as the supply lasts.

4. Agrees not to submit the full report to another journal of general circulation.

Specifications

Brief Report. The Brief Report should give a clear, condensed summary of the procedure of the study and as full an account of the results as space permits.

To insure that the Brief Report will be no longer than one printed page, its typescript, in-

Style of Manuscripts. Manuscripts must be prepared in the style described in the *Publication Manual of the American Psychological Association (1967 Revision)*, obtainable for \$1.50 from the Washington office of the APA.

Early Publication. Authors may secure the early publication of an acceptable article in the first issue to go to press after its acceptance by arranging for the article to be printed as extra pages, increasing the number of pages that subscribers receive. Information about charges for early publication may be obtained from the Managing Editor.

National Auxiliary Publications Service. To save printing costs, materials supplementary to articles may be deposited with the NAPS. To obtain copies, order by Document Number, as cited in the reference footnote, from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001. The indicated fee must be remitted in advance.

cluding all matter except the title and the author's lines, must not exceed 85 lines averaging 42 characters and spaces in length. Set the type-writer margins for short lines of 42 characters, which are 3.5 inches long in elite typing, and 4.2 inches long in pica.

The manuscript of the Brief Report must be *double spaced* throughout. Except for its short lines, it follows the standard style of the 1967 revision of the *APA Publication Manual*. Headings, tables, and references are avoided or, if essential, must be counted in the 85 lines. Each Brief Report must be accompanied by a footnote in the style below, which is typed on a separate sheet and *not* counted in the 85-line quota:

¹ An extended report of this study may be obtained without charge from John Doe (giving the author's full name and address) or for a fee from the American Society for Information Science. Order NAPS Document — from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 909 Third Avenue, New York, New York 10022; remitting \$ — for microfiche or \$ — for photocopies.

Extended report. Because the extended report is intended for photoduplication, and is not copy to be sent to a printer, its style should differ in several ways from that of other manuscripts: (a) The extended report should be typed with single spacing for economy in duplication. (b) Tables and figures should be placed adjacent to the text which refers to them. A caption should be typed below each figure. (c) Footnotes should be typed at the bottom of the page on which reference is made to them. In other respects, the full report is prepared in the style specified by the *Publication Manual*.

Abstracts. Manuscripts must be accompanied by an abstract of 100–120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or they may be obtained from the Editor or from the APA Central Office.

Copyrighted material. The following policies govern reprinting or other forms of reproduction of APA-copyrighted journal materials. Permission by the American Psychological Association must be requested for reprinting all tables, figures, and text of more than 500 words (total from one article) in length. There will be a charge of \$5.00 per page for such permission except when the requester is the author of the article to be reprinted, or the article is to be reproduced in a limited number of copies solely for instructional purposes.

Reprints. Authors receive 50 reprints gratis. Additional quantities may be ordered when returning proofs.

EFFECTS OF A MODEL AND INSTRUCTIONS ON GROUP VERBAL BEHAVIORS¹

CAROL WHALEN²

University of California, Los Angeles

This study assessed the relative efficacy of modeling and instructional approaches in increasing interpersonal openness in a group setting. Groups of four male college students participated in a leaderless discussion session after exposure to one of four manipulations: a film model of interpersonal openness plus detailed exhortative and descriptive instructions, film model plus minimal instructions, detailed instructions with no film, and minimal instructions only. It was found that Ss exposed to both the film model and the detailed instructions tended to engage in interpersonal openness, while groups in the other three conditions failed to do so, devoting most of their time to impersonal discussion. These findings were discussed in terms of the combined role of understanding and disinhibition in modifying verbal response classes.

In the past decade, modeling or imitation learning has received increasing emphasis as both a necessary and a real component of therapeutic behavior change. Authors using widely differing theoretical contexts have insisted that the psychotherapist serves as a model whose function is to demonstrate desirable behaviors so that his patients can learn through imitation (e.g., Bandura, 1965; Haley, 1963; Jourard, 1964; Masserman, 1957; Satir, 1964). The literature is also replete with evidence indicating that imitation learning not only should but actually does occur in traditional psychotherapeutic settings (e.g., Matarazzo, Wiens, & Saslow, 1965; Schefflen, 1963). After demonstrating that some form of modeling in psychotherapy is both inevitable and desirable, one must ask whether modeling techniques can be sys-

tematically and efficiently utilized to effect specific therapeutic goals.

Although a few early studies of modeling in therapylike situations can be found in the literature (e.g., Chittenden, 1942; Jones, 1924), most of the work has been done in the 1960s. Metz (1965) and Lovaas, Freitas, Nelson, and Whalen (1967) used a combination of imitation and reinforcement to teach schizophrenic children to perform a variety of self-care and social responses. Bandura, Grusec, and Menlove (1967) used a modeling approach to help nonpsychotic children overcome their fear of dogs. In their discussions of behavioral rehearsal techniques, Gittelman (1965) and Lazarus (1966) have described further uses of modeling to modify inappropriate or deficient behavioral repertoires.

Systematic attempts to change behavior through the provision of models often focus on verbal responses. For example, an imitation-reinforcement procedure has been used to reinstate simple verbal behaviors in mute psychotics (Sherman, 1965; Wilson & Walters, 1966) and to initiate speech in schizophrenic children (Lovaas, Berberich, Perloff, & Schaeffer, 1966).

There have also been a few experimental attempts to modify more complex aspects of verbal behavior through imitation procedures. Schwartz and Hawkins (1965), using patients as models and a therapist dispensing reinforcements, succeeded in either increasing or decreasing patient affect statements. Using

¹ This paper is based on a PhD dissertation submitted to the Department of Psychology, University of California, Los Angeles, 1967. The writer wishes to express her gratitude to the chairman of her doctoral committee, Tommy Tomlinson, and to Barbara Henker, Charles Nakamura, and Tom Trabasso for their generous help and support during the course of this investigation. Thanks are also due to Jill Waterman Janovici and Richard Whitney, without whose continued efforts the data collection phase of this investigation would not have been possible. This investigation was partially supported by United States Public Health Service Predoctoral Fellowship 5 F1-MH-20, 464-04.

² Requests for reprints should be sent to Carol Whalen, who is now at the Fairview State Hospital, Costa Mesa, California 92626. The author is also at the University of California, Irvine.

college student Ss, Marlatt, Jacobsen, Johnson, and Morrice (1966) modified the degree of problem admission through a similar modeling-reinforcement procedure. And finally, Truax and Carkhuff (1965) have described the use of a tape-recorded model to facilitate the learning by new patients of those verbal behaviors deemed desirable in a therapeutic encounter.

These studies demonstrate that imitation can involve the learning of entire classes of behaviors, such as verbal "affect" statements, as well as simple discrete responses, such as "That makes me angry." The goals of psychotherapy often include modification of such complex classes of verbal behaviors. For example, one of the major therapeutic objectives has traditionally been the development of insight or understanding. In the 1960s the focus has shifted away from the private, individual, and toward the social, communicative aspects of such insight. Bradford (1964), Jourard (1964), and Mowrer (1964) all affirm the importance of "interpersonal openness," the readiness to convey personal information about one's life, one's feelings, and one's impressions of other individuals.

The primary objective of the present investigation was to compare the efficacy of modeling and instructional approaches in manipulating complex verbal behaviors. As it appears that interpersonal openness is rapidly becoming a major goal of many behavior modification approaches, this response class was chosen as the major dependent variable.

METHOD

The study consisted of a 2×2 factorial design with film model (present or absent) and instructions (detailed or minimal) serving as the main variables. In the film-detailed instructions (F-DI) condition, Ss were presented with instructions describing interpersonal openness and directing them to engage in this type of interaction. In addition, these Ss viewed a film in which a group of individuals demonstrated or modeled the desired behaviors. This manipulation was seen as the condition of maximal influence, combining the informational and motivational cues provided by both models and instructions.

The Ss in the film-minimal instructions (F-MI) condition viewed the same film presented to F-DI Ss but received neither a description of interpersonal openness nor an exhortation to engage in such behavior. The intent of this manipulation was to provide evidence as to whether adults, like children,

will imitate complex and possibly embarrassing behaviors after merely observing models engage in such behaviors. In other words, can adult Ss discern what is expected of them and will they fulfill these expectations when a model is presented in the absence of exhortative or descriptive instructions?

In real-life settings verbal instructions of both a descriptive and a directive nature are often used alone to effect desired behavior change. In order to compare this procedure with the modeling approach, a third condition labeled no film-detailed instructions (NF-DI) was included. In this condition, Ss received instructions similar to those presented in the F-DI condition but were not exposed to the film model.

A fourth condition was included to provide a comparison level, delineating the verbal behaviors which occur in a group of individuals, meeting each other for the first time, when no definite structure is provided. The Ss in this fourth condition (NF-MI) received minimal instructions and no exposure to the film model.

On the basis of previous findings (Bandura, 1965) and pilot data, it was predicted that compared to the other three conditions, the F-DI manipulation would result in the greatest amount of interpersonal openness. Analogously, Ss in the F-MI, NF-DI, and NF-MI conditions were expected to engage in significantly more impersonal interactions than those exposed to the F-DI manipulation. No specific differences were predicted among the F-MI, NF-DI, and NF-MI conditions.

Subjects

The Ss were 128 male students who volunteered to participate in a group dynamics experiment in partial fulfillment of their introductory psychology course requirements. They participated in groups of four, each group consisting of members who had not previously met each other. The groups were randomly assigned to one of the four experimental conditions.

Film Model

The model consisted of a 12-minute, untitled, 16-millimeter film produced for this study by the Academic Communications Facility at the University of California, Los Angeles. A film was used in preference to live models so that all Ss would be exposed to an identical manipulation. Data from previous studies (e.g., Bandura, Ross, & Ross, 1963; Harby, 1962) indicate that film models are as potent as live models in eliciting imitative behavior.

In the experimental film a group of four male undergraduates conversed after implying that they had just met. Two of the film participants spent a few seconds conveying biographical information, and then a third member, interrupting these interchanges, maintained that the group members were not using the best method of getting to know each other. For the remainder of the time the participants talked on a more personal level, describing their anxieties and other feelings, and feeding back their

impressions—whether positive, negative, or neutral—to the other group members.

None of the Ss was acquainted with any of the four students in the film. Each film participant had previously attended and led sensitivity training sessions and thus was experienced in the types of interactions focused on in this study.

Attempts were made to make the film group appear as similar as possible in composition to the S groups (e.g., in age, sex, and dress). Similarity between Ss and models was maximized because of the possibility that large S-model differences might lead Ss to conclude that behavior matching was inappropriate, unnecessary, or impossible (Goldstein, Beller, & Sechrest, 1966; Hollander, 1958).

Instructions

Whether or not they were exposed to the film, all Ss received some form of verbal instruction. Instructions for each condition were tape recorded in order to control for possible intersession variability. A male E made the recordings. The detailed and minimal instructions follow.

Film-detailed instructions condition: This is a group dynamics laboratory. It is a rare opportunity to improve your ability to communicate with others by learning how you appear to them; that is, how you present yourself and what impressions they form about you. To accomplish this, it is important that you be open and honest with other members of the group and discuss freely your immediate feelings, attitudes, and opinions, both about yourself and others in the group. The information used to form your impressions of the other group members will come from their appearance, words, and behavior. We want you to feed back your impressions to the others as soon as the impression is formed in your mind. This will be difficult, but it is essential to obtain the maximum benefits from the experience. The amount you learn about yourself will depend on the extent to which you participate in the group.

To show you what we mean by open discussion of feelings and feedback of impressions, we are going to show you a film of a group similar in composition to yours. The group members in this film provide excellent examples of desirable types of group behavior. You will note that they are open in talking about themselves. In addition, they freely express their immediate impressions of other members of the group, and each member makes an honest effort to accept and examine the truth or falsity of the impression as it applies to him.

After the film you will form your own group, and we would like you to try and interact in a manner similar to that of the film group. No experimenter will participate in your group, but an observer will be watching through the one-way mirror in order to record the interaction variables we are examining.

After you have seen the film, you may begin. You will be notified when your time is up. Re-

member, it is important that you openly discuss your feelings and immediate impressions of others in the group, as well as invite similar comments about yourself.

Film-minimal instructions condition: This is a group dynamics laboratory. We are interested in examining some variables important in the development of group interaction. To enable us to study these variables we will ask you to interact as a group, but before you begin we are going to show you an example of what group interaction looks like. You will see subjects in a situation similar to the one in which you will find yourselves. After the film is over you may begin your own interaction. No experimenter will participate in your group, but an observer will be watching through the one-way mirror in order to record the interaction variables we are examining. You will be notified when your time is up.

Instructions for the two no-film conditions were identical to those for the respective film conditions except that all references to the film were deleted.

Procedure

The study was conducted in two adjoining rooms. The Ss entered a room containing four chairs placed around a circular table such that when seated all Ss were at least partially facing a one-way observation window. The E and raters recorded their observations from the adjoining room. The situation was arranged so that the observers both saw and heard the Ss, while the Ss heard nothing from the adjoining room and saw only a mirror.

The Ss were met by one of four male Es when they entered the experimental room and asked not to talk to the other group members before the experiment began. When the four Ss arrived, E played the tape-recorded instructions and, if appropriate, ran the film of the modeled interaction.

Following the experimental manipulation, E distributed 5 × 8 inch white index cards and a grease pen to Ss, requesting that each S write his first name on the card, fold it in half, and stand it up on the table in front of him so that the card was visible to the other Ss. The E then left the room and Ss began their conversation. After approximately 40 minutes had elapsed, E reentered the room to terminate the interchange, to request that Ss not discuss this experiment with their friends, and to thank them for their cooperation.

Dependent Variables

The dependent measures consisted of continuous ratings of five major verbal response classes. The classes were selected so as to include the behaviors modeled in the film as well as those which typically occur in newly formed groups. Four of the response classes were further broken down into component categories in order to allow a more precise analysis of the group verbal behaviors which occur in the various conditions. In all, 18 response categories were recorded.

The first two response classes, labeled "personal discussion" and "feedback," include the primary types of interaction modeled in the film. These two classes together comprise the more general response class labeled "interpersonal openness." The third response class, "impersonal discussion," is composed of verbal behaviors, such as conveying and requesting biographical information, which spontaneously occur quite frequently in newly formed groups. These three response classes were considered the primary foci of this investigation, since it was assumed that Ss who imitate or follow instructions would make utterances categorized as either personal discussion or feedback, while those who fail to imitate or follow instructions would engage in the conventional verbal behaviors categorized as impersonal discussion.

In order to obtain a comprehensive picture of the group interaction, two additional response classes were included in the behavioral analysis, although no specific predictions were advanced regarding their occurrence. One of these classes, labeled "group process," included all general discussions about the experimental manipulations and procedures. The remaining response class included several descriptive, noncontent aspects of communicative speech which frequently occur in a group setting.

All verbal responses which were either inaudible, unintelligible, or difficult to classify were included in a separate category which was labeled "unscorable utterances."

A more precise description of each response category follows.

1. *Personal discussion.* (a) *Personal self-disclosure:* Discussion of feelings, attitudes, and behaviors unrelated to current group interaction which are nonpublic, not ordinarily readily volunteered, and which may make the individual vulnerable to negative evaluations from others. Examples of actual comments categorized as personal self-disclosure are "I'm a status seeker like a lot of people," "I might even try homosexuality," or "I just broke down and started yelling and screaming." (b) *Immediate feelings:* Discussion of feelings and attitudes related to current group interaction which are relatively private, not ordinarily volunteered by most people, and which may make the individual vulnerable to negative evaluations from others (e.g., "I feel alienated here" or "Now I'm embarrassed"). (c) *Personal questions:* Questions relating to an individual's opinions, feelings, or actions which are usually considered to be private information: "Have you always been heavy?" or "I wondered how you felt when we were so silent."

2. *Feedback.* (a) *Positive feedback:* The conveying of an individual's positively valenced impressions of or reactions to the comments, appearance, or actions of another group member. This category includes compliments, flattery, etc. Examples are "I like the way you said that" or "You seem like an alert, intelligent guy." (b) *Negative feedback:* Negative evaluation of the personality, intelligence, actions, etc., of another group member. This category includes

hostility, ridicule, criticism, and implications that the individual is not telling the truth. Examples of comments rated as negative feedback are "You were a fool," "Big deal!" or "Your argument isn't very rational." (c) *Neutral feedback:* This category includes comments about a group member that cannot be evaluated as either positive or negative, as well as direct advice from one group member to another. Some examples are "You seem like the silent type" or "You ought to go up and tell her." (d) *Accept feedback:* Acceptance by a group member of the positive, negative, or neutral feedback he receives from other group members, for example, "Yeah, I guess you're right." (e) *Reject feedback:* Refusal to accept the feedback about oneself provided by other group members, for example, "No, you see it's really like this" (f) *Request feedback:* A direct request that the group or particular members provide S with an evaluation of himself, for example, "What do you think of me?"

3. *Impersonal discussion.* (a) *Impersonal self-disclosure:* Nonpersonal biographical information about oneself which is either generally accessible or readily volunteered in appropriate contexts by most people. Examples of this category include "I came from Chicago" or "I don't know the outskirts of the city well." (b) *Extragroup process:* Nonpersonal topics unrelated to the experiment or the group interactions—"All universities don't accept correspondence courses" or "I know a guy who is in the Navy now." (c) *Impersonal questions:* Questions regarding public, easily accessible, readily volunteered information about a group member—"Have you ever seen the San Diego campus?" or "How many units are you taking?"

4. *Group process.* Nonpersonal discussion relevant to the experiment such as comments regarding the experimental manipulations, the purposes of the experiment, the behavior of the Es or film participants, the delineation of a plan of action, etc. Examples: "Are they watching us?" "I guess they just want us to act normal" or "Let's all talk at the same time—that ought to goof them up."

5. *Descriptive aspects of communicative speech.* (a) *Agreement:* Any verbal indication of agreement, for example, "Uh huh" and "Yeah, that's right!" Head nods are not scored. (b) *Disagreement:* Any verbal indication of disagreement, for example, "NO" and "That's not true!" Head nods are not scored. (c) *Laughter:* Any amount of audible giggling or laughter by one or more group members. Smiles and grins are not scored. (d) *Silence:* This category is scored only if the duration of the silent period exceeds 3 seconds. (e) *Interruption:* Scored whenever one group member stops another from talking or when two or more individuals are speaking simultaneously.

6. *Unscorable utterances.* A scoring category for those remarks which are either inaudible, unintelligible, or too difficult to categorize immediately.

Rating Procedure and Apparatus

For 26 of the 32 groups, two trained raters independently categorized each utterance that occurred

during the discussion. Each rater used an Esterline-Angus event recorder and a 12-button operating panel. This apparatus provided a continuous record of both the frequency and the duration of each utterance in each category. The resultant scores describe total group interaction; the verbalizations of individual group members were not rated separately.

Analysis of Data

The frequency scores presented in the following analyses are actual frequency counts of each type of occurrence. The duration scores are not actual time periods. The exact number of seconds devoted to each response category can be obtained by multiplying each duration score by a factor of 7.5. In the statistical analyses the "unscorable utterances" were considered a separate category, thus bringing the total number of categories to 19.

There was some variation among groups in total interaction time. In order to make total scores comparable, the groups were equated by using the length of the shortest group as a cut-off criterion. Thus only data derived from the first 36 minutes of each group interaction were entered into the analyses.

Reliability. In order to assess the relationship between each rater's scores, a reliability analysis was performed on data from those groups for which two sets of ratings were obtained. Pearson product-moment correlation coefficients were computed from total scores for each group. The analysis was done separately according to condition, category, and type of score (frequency or duration), giving a possible total of $4 \times 19 \times 2$ or 152 analyses. However, in those cases in which there were fewer than two pairs of scores, the correlation coefficient was not computed since it would have no meaning. This criterion resulted in the elimination of 16 or approximately 10% of the possible correlation coefficients, leaving a total of 136 reliability measures.

After the correlation coefficients were computed, they were transformed, using Fisher's z coefficients, so that they could be averaged across conditions within each category. The z scores were weighted according to the number of groups in each condition, and then the arithmetic mean of the z 's across conditions was found. Finally, the mean z 's were converted into their corresponding correlations (Guilford, 1950). These final correlations were then used in evaluating the reliability of the rating system. A mean correlation coefficient of .75 was selected as the requisite minimum for acceptable reliability. Data from any category for which the mean correlation failed to reach this level were excluded from subsequent analyses.

Before transformation and averaging, the interrater reliability coefficients ranged from -.11 to 1.0. More than 60% of these coefficients reached .90, and 83% exceeded .74. After transformation and averaging, the correlation coefficients ranged from .10 to .99. Sixty-six percent were at or above .90, and 92% reached .75. Frequency data from the category "acceptance of feedback" and duration data from the two categories "group feelings" and "agreement"

were not considered. These data were omitted because interrater reliability did not reach the preestablished criterion of $r = .75$ for these three categories.

In summary, these data clearly indicate that the system of content analysis devised for and used in this investigation can be reliably applied by trained raters. For 15 of the 18 response categories, the interrater reliability indexes reached the preestablished criterion of $r = .75$ for both frequency and duration scores (see Whalen, 1967).

Mean proportions. In order to provide a descriptive comparison of the differential occurrences of each category in the four conditions and in the actual film, mean proportions were computed. For a given condition, the means for all reliable categories were first totaled, and then each individual category mean was divided by this total mean. This analysis was performed separately on the data from the film model as well as from each of the four experimental conditions. Separate analyses were done for each rater and for each type of score. Overall proportions for each response class were computed in a similar manner.

Statistical analyses of group differences. The effects of the film and instructions manipulations were more formally assessed using Winer's (1962) procedure for two-way analyses of variance with unequal cell frequencies. Separate analyses were performed for each rater, for each type of score (frequency and duration), and for each reliable category. The main reason for separately analyzing the data from each rater instead of including the rater variable as a third factor in the design was that the two raters did not always rate the same S groups.

Method of analysis. In order to facilitate these analyses, the frequency and duration data from the event recorders were coded into a digital form and punched onto computer input cards. Fortran programs were written for the reliability and factorial analyses. The data were processed on an IBM 1410.

RESULTS

Figures 1 and 2 illustrate the proportion of total verbalization frequency and duration (averaged across raters) devoted to each verbal response class in the actual film and by the groups participating in each of the four experimental conditions. They indicate that certain quantitative and qualitative aspects of group verbal behavior are strongly influenced by the joint presentation of instructions and a film model. It can be seen that F-DI groups engaged in more personal discussion and feedback than did S s in any of the other three conditions. The proportions of total verbalization frequency and duration devoted to these two response classes by S s in the F-DI condition approached the actual

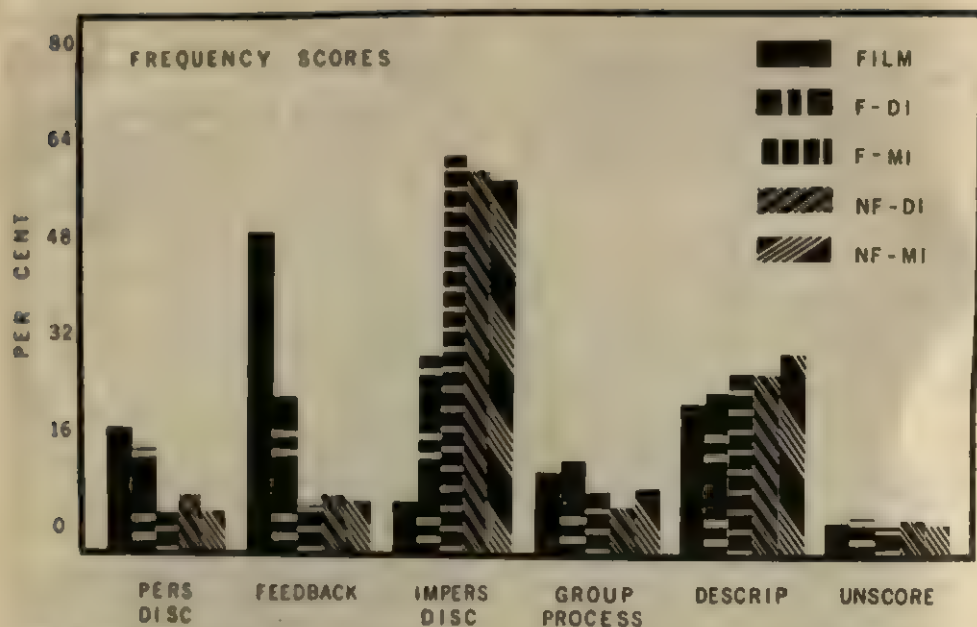


FIG. 1. Frequency scores: Relative proportion of total interaction devoted to each response class. (Code: Film = film model; F-DI = film-detailed instructions; F-MI = film-minimal instructions; NF-DI = no film-detailed instructions; NF-MI = no film-minimal instructions; PERS DISC = personal discussion; IMPERS DISC = impersonal discussion; DESCRIP = descriptive aspects; UNSCORE = unscorable utterances.)

film proportions, while the corresponding statistics for the remaining three conditions were markedly lower. These three latter conditions differed little among themselves in total time spent in personal discussion and feedback.

The proportion of total interaction time devoted to impersonal discussion by the F-DI groups also approached the analogous proportions for the actual film model. As can be seen from Figures 1 and 2, F-MI, NF-DI, and NF-MI groups engaged in markedly more impersonal verbalization than did the actual film group or Ss in the F-DI condition. Again, there was little variation in these scores among the F-MI, NF-DI, and NF-MI conditions.

In summary, relative to the other three conditions, the F-DI groups more closely matched the actual film model in the three response classes for which experimental predictions were advanced, namely, personal discussion, feedback, and impersonal discussion. In contrast, only minor differences were noted among the actual film and the four experi-

mental conditions regarding the proportion of total frequency and duration devoted to the remaining response classes, that is, group process and descriptive aspects, and to the unscorable utterances category.

Statistical Analysis of Film and Instructions Effects

Tables 1-4 present for each condition the mean frequency and duration scores ascribed to the individual categories by each of the two raters. These tables also summarize the tests of statistical significance. The following discussion is based on these means and the results of 12 separate *F* tests for each category: tests on both main effects (film and instructions) as well as the interaction for each rater and for each type of score.

Personal discussion. As shown in Table 1, the overall personal discussion frequencies and durations of the F-DI groups were two to eight times as large as those for groups in the other three conditions. Analyses of variance revealed a significant instructions variable in all but one case, regardless of

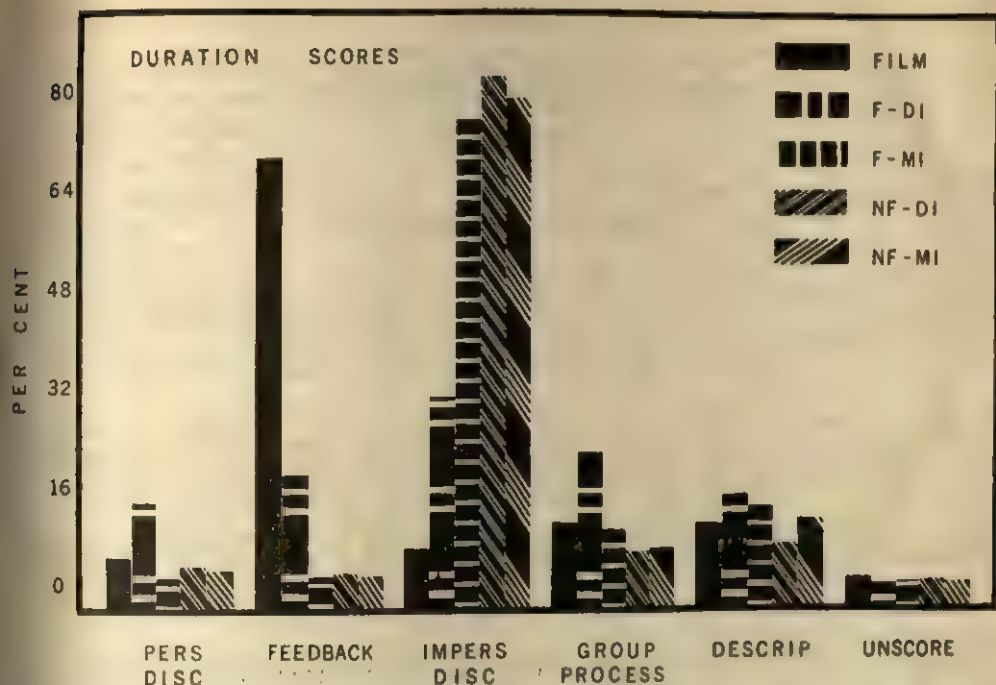


FIG. 2. Duration scores: Relative proportion of total interaction devoted to each response class. (Code: Film = film model; F-DI = film-detailed instructions; F-MI = film-minimal instructions; NF-DI = no film-detailed instructions; NF-MI = no film-minimal instructions; PERS DISC = personal discussion; IMPERS DISC = impersonal discussion; DESCRIP = descriptive aspects; UNSCORE = unscorable utterances.)

TABLE 1
MEAN FREQUENCY AND DURATION SCORES AND ANALYSES OF VARIANCE
FOR THE PERSONAL DISCUSSION RESPONSE CLASS

Category	Score	Rater	Condition means				Analyses of variance					
			F-DI	F-MI	NF-DI	NF-MI	Film		Instructions		F × I	
							F	p	F	p	F	p
Personal self-disclosure	Frequency	1	12.3	3.0	6.2	2.7	1.87	ns	7.65	.05	1.55	ns
		2	18.6	3.1	5.7	3.7	4.13	.10	8.35	.01	4.86	.05
	Duration	1	23.9	2.6	7.4	2.9	3.81	.10	9.58	.01	4.01	.10
		2	34.2	4.1	7.9	10.4	2.80	ns	5.34	.05	7.40	.05
Immediate feeling	Frequency	1	2.5	.1	.1	0.0	18.30	.01	18.30	.01	14.82	.01
		2	4.0	.4	0.0	0.0	8.61	.01	5.60	.05	5.60	.05
	Duration	1	Omitted because of unreliability									
		2										
Personal question	Frequency	1	4.2	1.1	2.0	.4	3.92	.10	10.17	.01	1.03	ns
		2	6.6	.6	1.0	.7	5.19	.05	6.93	.05	5.56	.05
	Duration	1	2.0	.4	.8	.1	7.22	.05	16.54	.01	1.94	ns
		2	2.5	.4	.4	.2	3.72	.10	4.00	.10	2.86	ns

Note.—F-DI = film-detailed instructions; F-MI = film-minimal instructions; NF-DI = no film-detailed instructions; NF-MI = no film-minimal instructions.

TABLE 2
MEAN FREQUENCY AND DURATION SCORES AND ANALYSES OF VARIANCE
FOR THE FEEDBACK RESPONSE CLASS

Category	Score	Rater	Condition means				Analyses of variance					
			F-DI	F-MI	NF-DI	NF-MI	Film		Instructions		F × I	
							<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>
Positive feedback	Frequency	1	8.7	.6	1.6	0.0	5.72	.05	9.09	.01	4.00	.10
		2	8.7	.7	.3	0.0	14.02	.01	11.33	.01	10.01	.01
	Duration	1	9.7	.3	1.1	0.0	4.68	.05	6.62	.05	4.04	.10
		2	9.3	.5	.2	0.0	8.08	.01	7.08	.05	6.57	.05
Negative feedback	Frequency	1	7.0	1.0	.4	.7	21.91	.01	14.70	.01	18.44	.01
		2	7.0	1.3	.9	.8	10.43	.01	17.99	.01	7.76	.01
	Duration	1	7.2	.6	.1	.6	10.38	.01	7.36	.05	10.32	.01
		2	9.2	1.3	.5	.5	9.62	.01	6.51	.05	6.76	.05
Neutral feedback	Frequency	1	17.7	3.6	4.6	4.0	8.70	.01	11.67	.01	9.76	.01
		2	18.9	4.9	6.7	7.0	3.80	.10	7.24	.05	7.77	.01
	Duration	1	18.4	2.1	3.4	3.3	9.18	.01	13.10	.01	12.66	.01
		2	18.6	3.1	5.2	5.3	4.33	.05	8.24	.01	8.45	.01
Request feedback	Frequency	1	3.0	.1	0.0	.1	16.73	.01	14.04	.01	17.13	.01
		2	2.6	0.0	0.0	.3	13.62	.01	13.62	.01	22.70	.01
	Duration	1	2.1	.1	0.0	.1	12.82	.01	11.29	.01	14.16	.01
		2	1.7	0.0	0.0	.2	12.50	.01	12.50	.01	18.69	.01
Accept feedback	Frequency	1	Omitted because of unreliability									
		2	Omitted because of unreliability									
	Duration	1	1.2	.1	.1	.1	13.59	.01	11.94	.01	12.32	.01
		2	2.1	.1	.1	.1	17.63	.01	19.63	.01	18.06	.01
Reject feedback	Frequency	1	1.5	.4	.1	0.0	9.59	.01	4.90	.05	3.13	.10
		2	1.9	.3	.4	.5	3.12	.10	4.04	.10	5.54	.05
	Duration	1	1.1	.1	.1	0.0	9.83	.01	7.07	.05	5.38	.05
		2	1.1	.1	.1	.3	3.71	.10	3.85	.10	7.24	.05

Note.—F-DI = film-detailed instructions; F-MI = film-minimal instructions; NF-DI = no film-detailed instructions; NF-MI = no film-minimal instructions.

category, score, or rater. The one exception was that for the personal question duration scores of Rater 2, differences in the instructions dimension only approached significance. The film variable was significant only for group feelings frequency scores of both raters, personal question frequency scores of Rater 2, and personal question duration scores of Rater 1. Rater 2's frequency and duration scores for the personal self-disclosure category led to a significant interaction, as did both raters' frequency scores for the group feelings category. The only other significant interaction effect was found for Rater 2's personal question frequency scores.

Feedback. As with the personal discussion categories, feedback occurred markedly more often in the F-DI condition than in any of the three remaining conditions (see Table 2). The F-DI Ss engaged in feedback as much as eight times as often and for 12 times as long as did Ss in the F-MI, NF-DI, and NF-MI groups.

For the film variable, 19 of the 22 *F* tests were significant, and the remaining three approached significance. Similarly, 20 of the 22 *F* tests for the instructions variable and 19 of the 22 *F* tests for the interaction were significant, while the results of the remaining tests approached significance.

TABLE 3
MEAN FREQUENCY AND DURATION SCORES AND ANALYSES OF VARIANCE
FOR THE IMPERSONAL DISCUSSION RESPONSE CLASS

Category	Score	Rater	Condition means				Analyses of variance					
			F-DI	F-MI	NF-DI	NF-MI	Film		Instructions		F × I	
							F	p	F	p	F	p
Impersonal self-disclosure	Frequency	1	20.8	44.1	32.2	35.6	.08	ns	6.72	.05	3.78	.10
		2	25.9	46.3	33.4	34.7	.16	ns	4.53	.05	3.51	.10
	Duration	1	25.7	62.5	62.0	49.3	1.22	ns	1.33	ns	5.60	.05
		2	32.9	67.8	65.4	49.5	.30	ns	.53	ns	3.80	.10
Extragroup process	Frequency	1	13.3	39.2	33.9	43.0	8.42	.01	17.52	.01	4.02	.10
		2	18.5	38.0	35.0	38.8	5.22	.05	9.46	.01	4.26	.05
	Duration	1	32.2	142.6	153.7	155.0	19.15	.01	13.34	.01	12.72	.01
		2	42.7	138.8	151.1	160.9	18.60	.01	12.25	.01	8.13	.01
Impersonal question	Frequency	1	11.2	24.6	17.6	11.3	1.02	ns	1.09	ns	8.46	.01
		2	11.5	25.3	17.1	13.0	1.12	ns	2.35	ns	8.07	.01
	Duration	1	3.5	7.7	5.7	3.2	1.08	ns	.55	ns	9.07	.01
		2	3.0	6.5	4.4	3.6	.76	ns	2.43	ns	6.29	.05

Note.—F-DI = film detailed instructions; F-MI = film-minimal instructions; NF-DI = no film-detailed instructions; NF-MI = no film-minimal instructions.

Impersonal discussion. As can be seen from Table 3, means for this response class were generally the reverse of those for the personal discussion and feedback categories. Relative to the F-DI groups, F-MI, NF-DI, and NF-MI Ss engaged in impersonal discussion about twice as often and for about three times as long. Most of the comments made by Ss in the F-MI, NF-DI, and NF-MI conditions were classified as impersonal discussion, while fewer than a third of the utterances of F-DI Ss were scored in this response class.

Differences due to the film variable reached a significant level only for the extragroup process category (for both types of scores and both raters). The instructions variable was also significant for all four tests in the extragroup process category. In addition, the impersonal self-disclosure frequency scores for both raters were significantly different on the instructions dimension. Tests of the interaction revealed a significant effect for the impersonal question category regardless of rater and score. The interaction effect for

TABLE 4
MEAN FREQUENCY AND DURATION SCORES AND ANALYSES OF VARIANCE
FOR THE GROUP PROCESS RESPONSE CLASS

Category	Score	Rater	Condition means				Analyses of variance					
			F-DI	F-MI	NF-DI	NF-MI	Film		Instructions		F × I	
							F	p	F	p	F	p
Group process	Frequency	1	20.8	10.7	6.6	9.9	4.70	.05	.97	ns	3.65	.10
		2	18.9	11.0	4.4	9.0	6.90	.05	.27	ns	3.96	.10
	Duration	1	49.3	24.6	19.1	16.1	4.68	.05	2.42	ns	1.48	ns
		2	47.0	24.9	11.3	14.7	11.08	.01	1.83	ns	3.42	.10

Note.—F-DI = film-detailed instructions; F-MI = film-minimal instructions; NF-DI = no film-detailed instructions; NF-MI = no film-minimal instructions.

the extragroup process category was also significant for Rater 2's frequency scores and the duration scores of both raters. All tests of the interaction which did not reach a significant level did approach significance.

Group process. It can be seen from Table 4 that F-DI Ss engaged in somewhat greater amounts of group process verbalization than did Ss in the other three conditions. The film model variable was significant for the frequency and duration scores of both raters; the interaction analysis of frequency scores from each rater and duration scores from Rater 2 approached significance.

Descriptive aspects. This response class includes about 20-30% of all interactions in each condition. The mean frequency and duration scores for the four conditions are generally quite similar. Only one of the 54 analyses of the data from the five categories comprising this response class reached a significant level (Rater 1's frequency scores for the agreement category), indicating that neither the film nor the instructions influenced this response class.

Unscoreable utterances. According to the scores of Rater 1, this relatively rare category occurred more often and for a longer period of time in the F-DI condition than in the other three conditions. Significant differences were found in the instructions analysis of Rater 1's frequency and duration data. None of the other analyses of data from either rater revealed significant differences in this category. (See Whalen, 1967, for the means and F values for the descriptive and unscoreable response classes.)

In summary, for all 12 categories comprising the three critical response classes (personal discussion, feedback, and impersonal discussion), some of the F tests reached significance. However, the categories varied as to whether the significant effect rested in the instructions, the film, the interaction, or some combination of these three factors. As previously pointed out and illustrated in the tables, the major group differences appear primarily due to the F-DI condition alone. Compared to the relative magnitudes of F-DI scores, differences among scores in the other three conditions seem unimportant.

DISCUSSION

The data obtained from this experiment indicate that exposure to a filmed model of interpersonal openness, when preceded by detailed exhortative and descriptive instructions, facilitates the expression of interpersonal openness and inhibits impersonal discussion in the observers. When presented without the other, neither the film nor the detailed instructions has these facilitative and inhibitive effects. In fact, the performances of Ss who were exposed to the instructions with no film and to the film with only minimal instructions were markedly similar to those of control Ss who were exposed to neither the film nor the detailed instructions.

In addition to the predicted effects, it was found that F-DI as well as F-MI Ss engaged in more group process than did Ss in the remaining two conditions (see Table 4). It is likely that this effect resulted partially from the novel and unexpected stimuli provided by the film, leading Ss exposed to the film to spend more time discussing the experimental manipulations than Ss receiving only the more traditional tape-recorded instructions. Stated simply, Ss viewing the film had more to talk about regarding the experiment.

It is interesting, however, that the group process engaged in by Ss in the two film conditions differed in quality. It appeared that F-MI groups spent significant portions of time criticizing the film participants and justifying their own failure to imitate, while the F-DI Ss seemed to be concerned with checking their own behavior against that of the film participants in order to see how close a match they were achieving. Unfortunately, no objective measures of these two types of group process were taken.

According to accepted definitions of imitation (e.g., Bandura & Walters, 1963), it can be said that F-DI Ss imitated—they matched the behavior of a model. The fact that F-MI Ss failed to imitate raises many interesting questions. A survey of the imitation literature (e.g., Bandura et al., 1963) reveals that experimental manipulations similar to those of the F-MI condition are usually sufficient to elicit modeling. One can use either a

motivational or a cognitive framework, or a combination of the two, for explaining the differences between F-DI and F-MI Ss.

Motivational Framework: The Inhibition Hypothesis

There are three alternative hypotheses that could be advanced within a motivational framework: (a) The act of imitation itself may have been inhibited; Ss could have felt that they would be cheating or unmanly if they matched the demonstrated behaviors. (b) F-MI Ss may have failed to engage in interpersonal openness because impersonal discussion is the customary verbal behavior in a new group, and the Ss had no reason to exert the extra energy needed to change these behaviors ("Law of Least Effort"). (c) The expression of interpersonal openness may have been perceived as risky and potentially embarrassing.

The first two explanations, although feasible, seem unlikely on both logical and empirical grounds (see Whalen, 1967, for a further discussion of these notions). The embarrassment explanation appears more plausible. The assumption is that interpersonal openness is often perceived as risky because it depicts the individual as weak and inadequate (Lakin & Carson, 1966). Because of this perceived risk or fear of punishment, the expression of interpersonal openness is inhibited. One could postulate that both the detailed instructions and the film model served to partially disinhibit interpersonal openness. The instructions depicted this behavior as highly desirable; they suggested that interpersonal openness was not only permissible, but also expected in this setting. The film further disinhibited Ss by demonstrating that the consequences of engaging in self-disclosure and feedback are positive rather than negative. It could be that, while neither of these two "disinhibitors" was itself potent enough to counteract the fear of punishment, the combined effect of the two was sufficient for this purpose.

The F-MI Ss, lacking the external inducement provided by detailed instructions, apparently created their own standard of behavior within the experimental session. Indeed, it appeared that in the F-MI groups, Ss often

spent some time criticizing the behavior of the film participants and developing arguments contraindicating the expression of interpersonal openness. These conversations, which generally occurred early in the group session, appeared to be attempts on Ss' parts to convince themselves and each other that imitative behavior in this instance was inadvisable. This speculation regarding explicit and implicit inducements is consistent with Berelson and Steiner's (1964) conclusion that the less definite the standards external to the group, the more control the group itself can exercise.

Cognitive Framework: The Ignorance Hypothesis

Rather than focusing on the motivational aspects of imitation, one could view the results of this study within an attention-information context. Thibaut and Kelley (1959), for example, have detailed a norm-sending process consisting of three components: rule statements, surveillance-evaluation, and application of sanctions (rewards and punishments). In the present study the film-instructions manipulation could be considered the rule statement; the presence of the other group members and the knowledge that the Es were observing from the adjoining room might serve an implicit surveillance-evaluation function; and the other group members as well as S himself could apply the sanctions.

This conceptualization bears a marked resemblance to a simple learning-reinforcement framework. Perhaps the primary difference is that it focuses on the learning of rules or strategies rather than discrete acts. As previously stated, it stresses the cognitive as well as the overt motor components of behavior. Although such a focus is not new in the study of imitation (Asch, 1952; Koffka, 1925), little emphasis has been placed on problems of attention and understanding in the recent research on modeling. Perhaps cognition has been ignored because data derived from the tasks typically employed in imitation experiments can be more parsimoniously explained in simple stimulus-response terms. A weighty cognitive framework seems unnecessary when describing dis-

crete motor acts or two-choice decisions, since Ss imitate these simple behaviors even when problems of attention and understanding are ignored in the experimental manipulations (e.g., Kimbrell & Blake, 1958).

The present experiment differed from the usual imitation study in that complex verbal response classes rather than simple discrete acts were modeled; successful imitation involved matching general verbal categories rather than simply mimicking the exact behavior observed in the film model. Perhaps in this latter situation some attention must be paid to orienting S toward the relevant variables if imitation is to occur.

It is interesting to note that this "cognitive" view of behavior change has recently been advocated by numerous experimental as well as clinical psychologists (e.g., Breger & McGaugh, 1965) and has formed the basis of several systems of psychotherapy (e.g., Haley, 1963). Moreover, despite a general disregard for cognitive components in empirical studies of imitation, recent studies and theories of modeling behavior (e.g., Bandura, 1965) reveal an incipient trend toward considering these aspects.

In summary, it would appear that the norm-sending notion (which stresses the importance of attention and information) and the disinhibition notion (which assumes interpersonal openness is viewed as potentially embarrassing or dangerous) provide feasible frameworks for discussing the present findings.

REFERENCES

- ASCH, S. E. *Social psychology*. Englewood Cliffs, N. J.: Prentice-Hall, 1952.
- BANDURA, A. Vicarious processes: A case of no-trial learning. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. Vol. 2. New York: Academic Press, 1965.
- BANDURA, A., GRUSEC, J. E., & MENLOVE, F. L. Vicarious extinction of avoidance behavior. *Journal of Personality and Social Psychology*, 1967, 5, 16-23.
- BANDURA, A., ROSS, D., & ROSS, S. A. Imitation of film-mediated aggressive models. *Journal of Abnormal and Social Psychology*, 1963, 66, 3-11.
- BANDURA, A., & WALTERS, R. H. *Social learning and personality development*. New York: Holt, Rinehart & Winston, 1963.
- BERELSON, B., & STEINER, G. A. *Human behavior: An inventory of scientific findings*. New York: Harcourt, Brace & World, 1964.
- BRADFORD, L. P. Membership and the learning process. In L. P. Bradford, J. R. Gibb, & K. D. Benne (Eds.), *T-group theory and laboratory method*. New York: Wiley, 1964.
- BREGER, L., & MCGAUGH, J. L. Critique and reformulation of "learning-theory" approaches to psychotherapy and neurosis. *Psychological Bulletin*, 1965, 63, 338-358.
- CHITTENDEN, G. E. An experimental study in measuring and modifying assertive behavior in young children. *Monographs of the Society for Research in Child Development*, 1942, 7, No. 1 (Serial No. 31).
- GITTELMAN, M. Behavior rehearsal as a technique in child treatment. *Journal of Child Psychology and Psychiatry*, 1965, 6, 251-255.
- GOLDSTEIN, A. P., HELLER, K., & SECHREST, L. B. *Psychotherapy and the psychology of behavior change*. New York: Wiley, 1966.
- GUILFORD, J. P. *Fundamental statistics in psychology and education*. New York: McGraw-Hill, 1950.
- HALEY, J. *Strategies of psychotherapy*. New York: Grune & Stratton, 1963.
- HARBY, S. F. Evaluation of a procedure for using daylight projection of film loops in teaching skills. Unpublished manuscript, Pennsylvania State College Instructional Film Research Program, 1962.
- HOLLANDER, E. P. Conformity, status and idiosyncrasy credit. *Psychological Review*, 1958, 65, 117-127.
- JONES, M. C. The elimination of children's fears. *Journal of Experimental Psychology*, 1924, 7, 383-390.
- JOURARD, S. M. *The transparent self*. Princeton: Van Nostrand, 1964.
- KIMBRELL, D. L., & BLAKE, R. R. Motivational factors in violation of a prohibition. *Journal of Abnormal and Social Psychology*, 1958, 56, 132-133.
- KOFFKA, K. *The growth of the mind*. New York: Harcourt, Brace, 1925.
- LAKIN, M., & CARSON, R. C. A therapeutic vehicle in search of a theory of therapy. *Journal of Applied Behavioral Science*, 1966, 2, 27-40.
- LAZARUS, A. A. Behaviour rehearsal vs. non-directive therapy vs. advice in effecting behaviour change. *Behavior Research and Therapy*, 1966, 4, 209-212.
- LOVAAS, O. I., BERBERICH, J. P., PERLOFF, B. F., & SCHAEFFER, B. Acquisition of imitative speech by schizophrenic children. *Science*, 1966, 151, 705-707.
- LOVAAS, O. I., FREITAS, L., NELSON, K., & WHALEN, C. K. The establishment of imitation and its use for the development of complex behaviors in schizophrenic children. *Behavior Research and Therapy*, 1967, 5, 171-181.
- MARLATT, G. A., JACOBSEN, E. A., JOHNSON, D. L., & MORRICE, D. J. Effect of exposure to a model receiving varied informational feedback upon consequent behavior in an interview. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, May 1966.

- MASSERMAN, J. H. Evolution vs. "revolution" in psychotherapy: A biodynamic integration. *Behavioral Science*, 1957, 2, 89-100.
- MATARAZZO, J. D., WIENS, A. N., & SASLOW, G. Studies of interview speech behavior. In L. Krasner & L. P. Ullman (Eds.), *Research in behavior modification*. New York: Holt, Rinehart & Winston, 1965.
- METZ, J. R. Conditioning generalized imitation in autistic children. *Journal of Experimental Child Psychology*, 1965, 2, 389-399.
- MOWRER, O. H. The new group therapy. New York: Van Nostrand, 1964.
- SATIR, V. *Conjoint family therapy: A guide to theory and technique*. Palo Alto: Science & Behavior Books, 1964.
- SCHEFLEN, A. E. Communication and regulation in psychotherapy. *Psychiatry*, 1963, 26, 126-136.
- SCHWARTZ, A. N., & HAWKINS, H. L. Patient models and affect statements in group therapy. In, *Proceedings of the 73rd annual convention of the American Psychological Association*. Washington, D. C.: American Psychological Association, 1965.
- SHERMAN, J. A. Use of reinforcement and imitation to reinstate verbal behavior in mute psychotics. *Journal of Abnormal Psychology*, 1965, 70, 155-164.
- THIBAUT, J. W., & KELLEY, H. H. *The social psychology of groups*. New York: Wiley, 1959.
- TRUAX, C. B., & CARKHUFF, R. R. Personality change in hospitalized mental patients during group psychotherapy as a function of the use of alternate sessions and vicarious therapy pre-training. *Journal of Clinical Psychology*, 1965, 21, 225-228.
- WHALEN, C. K. The effects of a model and instructions on group verbal behaviors. Unpublished doctoral dissertation, University of California, Los Angeles, 1967.
- WILSON, F. S., & WALTERS, R. H. Modification of speech output of near-mute schizophrenics through social-learning procedures. *Behavior Research and Therapy*, 1966, 4, 59-67.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

(Received September 20, 1968)

EDITOR'S NOTE

The following discussion by Kenneth Heller was prepared at the request of the Editor.

EFFECTS OF MODELING PROCEDURES IN HELPING RELATIONSHIPS

KENNETH HELLER¹

Indiana University

The clinical utility of modeling techniques is discussed in terms of a growing trend in which psychotherapy is viewed as part of a broader psychology of behavior change. Modeling techniques can be important facilitators of change, but their application to clinically helpful relationships is impeded by a paucity of research investigating the operation of modeling procedures in clinical-like settings. Intermediate steps or "bridging" research is needed before experimentally derived hypotheses can be applied clinically. The Whalen study, which is one such "bridge," describes the effort needed for modeling procedures to be effective in adult groups and illustrates the need for constant attention to setting and context factors if a systematic psychology of behavior change is to be developed.

Throughout the history of psychotherapy, the major source of ideas about its conduct has been reports of experiences of practicing therapists. This is as true today as it was some 70 years ago when Breuer and Freud (1937) developed their theoretical notions about psychopathology derived from experiences with five cases of hysteria.

Most recently, the experimental laboratory has become a serious source of competing hypotheses about the nature of therapeutic interventions. Psychologists in seemingly divergent areas of experimental research have been trying, for some time, to understand the factors responsible for behavior change. The literature on the conditions that facilitate psychological change has been growing in such fields as learning, education, social influence, attitude change, persuasion, and cognitive restructuring. Clinical psychologists trained in academic settings have begun to wonder whether changes in therapeutic technique might be possible if, taking advantage of the experimental literature, psychotherapy became part of a broader psychology of behavior change.

The work of Whalen (1969) is one attempt to test, in a clinical-like setting, hypotheses derived from the experimental literature on observational learning and to suggest a clinical utility to teaching procedures. In discussing the Whalen

study, several points will be covered: the current research evidence concerned with the effects of models in personal interviews; models as facilitators of change in helping relationships; the limits of modeling procedures; and the function of laboratory analogue research in the acquisition of knowledge about behavior change in clinical settings.

Modeling Techniques in Personal Interviews

There have been a number of studies demonstrating the facilitating effects of modeling procedures on the discussion of personally revealing topics in interviews. Several of these studies, conducted at Indiana University, have been described elsewhere (Heller, 1968; Heller & Marlatt, 1969). In summary, this research has found that modeling procedures provide an effective means of teaching interview role behaviors, particularly when task instructions are ambiguous (Marlatt, 1968b). Modeling procedures were found to be more effective than verbal reinforcement in eliciting and sustaining self-disclosure (Marlatt, 1968a) and were also more effective when consequences to the model were either positive or neutral (Marlatt, Jacobson, Johnson, & Morrice, 1966). Teaching role expectations through modeling produces greater interview participation for chronic hospitalized psychotics as well as for more typical college undergraduates (Zerfas, 1965). However, for college student volunteers, instructions to "talk about your problems" without models can also produce high levels of self-disclosure in individual interviews (Jacobson, 1968).

The study by Whalen (1969) found that in a group setting neither instructions nor models alone could produce high levels of self-disclosure.

¹ At the time of writing this review, the author was on Sabbatical leave at the Laboratory of Community Psychiatry, Harvard Medical School, and was supported in part by Special Research Fellowship MH-40,558 from the National Institute of Mental Health.

Requests for reprints should be sent to Kenneth Heller, Department of Psychology, Indiana University, Bloomington, Indiana 47401.

A combination of detailed, exhortative instructions plus modeling were required. Her results point to the importance of context and setting factors in research of this type. The meaningful discussion of personal concerns and fears is often perceived as potentially risky and embarrassing. Within the context of an individual interview, either models or detailed instructions appear sufficient to elicit self-disclosure. In a group setting in which *S* faces complete strangers, the risk of appearing weak and inadequate is greater, and personal openness is inhibited. In the Whalen study, neither modeling nor instructions were, by themselves, capable of removing that inhibition. This study reminds us that therapeutic tools are not universal panaceas, indiscriminately effective under all circumstances.

Models as Facilitators of Change in Helping Relationships

Since observational learning is so commonplace in normal socialization, it is a wonder that psychotherapists did not use it earlier as part of their therapeutic procedures. Undoubtedly, therapist ideology had an important role in its exclusion. The model of patient psychopathology with which therapists traditionally oriented their work primarily involved emotional and cognitive inhibitions. Even today, most of our current therapies assume that what went wrong initially in the patient's life is that anxiety engendered by the disapproval of others has blocked self-expression. Inhibitions lead to anxiety which in turn provides fuel for further inhibition. The patient becomes caught in a destructive cycle which leaves him unable to think, feel, or act appropriately.

The inhibition model of pathology predicts that attempts to foster observational learning will fail because the patient's ability to perform adequately is blocked by anxiety. Indeed, there is anecdotal evidence to substantiate the belief that fearful patients will be unresponsive to model behavior that is too "healthy" or bold. For example, Bandura, Grusec, and Menlove (1967) studied the effects of modeling procedures in reducing fear of dogs in nursery school children. Pilot work revealed that when a peer model approached the dog boldly from the beginning, and displayed an early close interaction with the dog, the fearful children actively avoided looking at the model's performance and were reluctant to participate in subsequent sessions.

With highly fearful patients, the therapist has more than one alternative available to reduce that fear. He can adopt the more typical pro-

cedure of working toward the extinction of fears by first removing verbal inhibitions, on the assumption that indirect, symbolic disinhibition of fears plus verbal understanding of them will lead to overt behavior change. However, the therapist can also deal with the patient's fears by encouraging approach responses more directly. The Bandura et al. (1967) study illustrates this point. Dog-phobic nursery school children were exposed to one of four treatment conditions. One group observed a peer model exhibit progressively stronger approach responses to a dog within the positive context of a party. A second group observed the model's behavior in a neutral context. A third group observed the dog during a party with the peer model absent, while a fourth group participated in a party with both the model and the dog absent. Results indicated that *Ss* who observed the model interact nonanxiously with the dog in either a positive or a neutral context displayed significantly greater approach responses to the dog, which persisted through a 1-month follow-up. A further extension by Bandura and Menlove (1968) revealed that filmed models interacting with dogs can be used in place of live models. Greatest effectiveness for the filmed models is achieved by a broad sampling of models and aversive stimuli. In other words, observation of different models varying in age, as they interact with various dogs, produces greater approach behavior in fearful children than observation of single-model single-dog interaction.

When the patient's primary problem does not involve areas of neurotic inhibition, verbal insight-oriented therapy alone may be insufficient. For example, many patients do not possess within their response repertoires the behaviors necessary for successful therapy or for productive real-life behavior. Limited experience makes them appear unsuitable for therapy; they need opportunities to learn prosocial behavior. If a therapist believes that he has a role in maximizing behavior change for patients with impoverished behavioral repertoires, he might wish to turn to the literature on imitation learning in his search for new therapeutic procedures (see Goldstein, Heller, & Sechrest, 1966, pp. 197-201 and 261-275).

As a therapeutic tool, modeling procedures can be used to weaken or remove inhibitory responses in the observer, to aid in teaching new behaviors, to help establish the timing of responses, and to provide a means of learning inhibitory controls. With the current emphasis on community mental health, psychologists are being called upon to serve in a variety of helping relationships besides psychotherapy. In these

new settings, similar uses of modeling procedures might be expected. For example, some consultants teach interviewing techniques to consultees (e.g., nurses, case aides, para-professionals) by having the latter observe the work of more experienced professionals. In a manner similar to that used by Whalen (1969), filmed and taped examples of encounter groups have been used to stimulate interpersonal openness in sensitivity training groups. We are now at a stage in which the potential utility of modeling procedures in psychologically helpful relationships is just beginning to be realized.

Limits of Modeling as a Therapeutic Technique

The mental health fields are pervaded by a philosophy of "oversell," in which techniques are put forth as panaceas for all types of problems. Inevitably, what follows each new fad is disillusionment resulting from inappropriate, over-generalized application. With reference to modeling techniques, faddishness might be retarded if we can clearly specify what we know about the conditions that affect the strength of observational learning. One way of beginning would be to specify why those in need of psychological help have not been able to profit from observing the behavior of others before entering psychotherapy. If observational learning is so widespread in our society, why are patients unable to resolve their difficulties on their own by observing the adequate solutions of others?

For some individuals, appropriate pro-social models are not available. For example, delinquents raised in a deviant subculture may have few opportunities to observe socially approved patterns of adjustment. Other individuals may be too inhibited to copy behaviors they see performed successfully, while still others may be too withdrawn to notice those behaviors. Included in this latter group are those chronically disturbed persons whose defensive style of reacting to stress removes them from normal social interaction. They illustrate the secondary effects of early emotional disturbance, namely social impoverishment associated with lack of experience.

For the groups cited above, either environmental circumstance or their own pathology prevents them from taking full advantage of the adaptive experiences of others. It might be expected that if they were provided with opportunities for observational learning in a benign therapeutic environment, these deficiencies might be overcome. The problem is that our knowledge of the factors facilitating observational learning in therapeutic settings is scant. Most of the re-

search on observational learning deals with socially acceptable behaviors and a good share of this work uses children as Ss. What is needed is more research examining the conditions necessary for adults to copy behavior which for them is personally unacceptable, inhibited, or socially disapproved. The Whalen (1969) study of self-disclosure in a group setting is one such attempt and reveals the difficulties in overcoming societal sanctions through modeling procedures alone.

Laboratory Research and Clinical Application

The transposition of the findings of laboratory research to areas of application quite different from its original setting is not accomplished without difficulty (Heller & Marlatt, 1969). For example, Hilgard and Bower (1966) describe six types of research which should be undertaken before principles of learning developed in the experimental laboratory can be adapted to school classrooms. Hovland (1959) and McGuire (1968), in social psychology, have also called for a more careful analysis of person and situation variables to resolve contradictory findings in studies of social influence and persuasibility. In the field of psychotherapy, Ford and Urban (1967) describe a three-stage process required to translate basic principles of behavior change to therapeutic practice. The point is that one should not expect to move directly from theoretical hypothesis to clinical application without prior information concerning the interaction of mediating variables such as personality, task, and setting. "Bridging" steps are required between hypothesis and application.

As illustrated in Figure 1, therapeutic hypotheses may be derived from clinical experience (their traditional source) or from behavior change research in other areas of psychology. Regardless of their source, factors thought to be responsible for change should be identified and studied in controlled laboratory settings that approximate clinical interactions. The purpose of clinical laboratory research is to determine what factors produce change, under what conditions they operate best, and how they should be combined to produce an effective therapeutic package. The therapeutic agents thus identified can then be studied in clinical field research to obtain information about the interaction of therapeutic ingredients with personality and setting characteristics that are part of actual treatment. Finally, the practitioner is in a position to apply this knowledge to individual patients with some degree of confidence in the expected outcome. Also, he can use the information gained

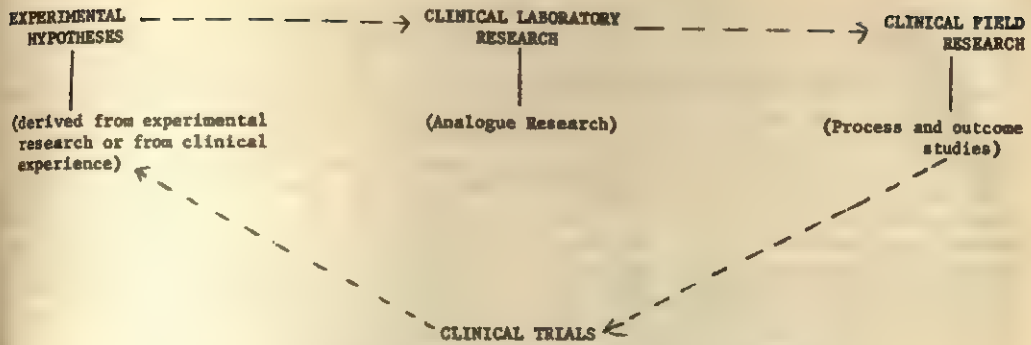


FIG. 1. The research steps needed between the discovery of psychological change factors and their clinical application.

from practice to develop further hypotheses about the relevant factors in the change process.

Needless to say, the above description does not represent the usual manner in which research in psychotherapy is conducted. Clinicians most often move directly from clinical hypotheses to clinical trials. Little attention is given to determining the exact factors that produce change in complex therapeutic procedures, or to the conditions under which each therapeutic ingredient is optimized. Therapeutic strategies are often expected to apply regardless of patient, therapist, or setting characteristics (Kiesler, 1966). Unfortunately, only after a therapeutic approach has become entrenched as a "school" with disciples committed to its perpetuation, do researchers begin to question the value of some of its therapeutic ingredients. An illustration of this point can be seen in recent developments in behavior therapy. After years of application and professional growth, the crucial ingredients in systematic desensitization are only now being given careful scrutiny.

The sequence in which the research steps between hypothesis and application are carried out is not at issue. What is important is that all are accomplished. With regard to the clinical application of modeling techniques, the necessary research is just beginning to accumulate. Clinical field studies have been conducted in which modeling procedures have been used to (a) teach psychotic patients and institutionalized delinquents what to expect from therapy (Truax, Wargo, Carkhuff, Kodman, & Moles, 1966; Zervas, 1965); (b) prepare unemployed juveniles for job application interviews (Sarason & Ganzer, in press); and (c) teach parents how to interact with their emotionally disturbed children (Guerney, 1964). Studies investigating the conditions under which modeling effects are optimized

in clinical-like laboratory settings are less well represented in the research literature. The Marlatt (1968a, 1968b), Marlatt et al. (1966), Jacobson (1968), and Whalen (1969) studies stand as rare exceptions.

REFERENCES

- BANDURA, A., GRUSEC, J. E., & MENLOVE, F. L. Vicarious extinction of avoidance behavior. *Journal of Personality and Social Psychology*, 1967, 5, 16-23.
- BANDURA, A., & MENLOVE, F. L. Factors determining vicarious extinction of avoidance behavior through symbolic modeling. *Journal of Personality and Social Psychology*, 1968, 8, 99-108.
- BREUER, J., & FREUD, S. *Studies in hysteria*. Boston: Beacon Press, 1937.
- FORD, D. H., & URBAN, H. B. Psychotherapy. *Annual Review of Psychology*, 1967, 18, 333-372.
- GOLDSTEIN, A. P., HELLER, K., & SECHREST, L. B. *Psychotherapy and the psychology of behavior change*. New York: Wiley, 1966.
- GUERNEY, B., JR. Filial therapy: Description and rationale. *Journal of Consulting Psychology*, 1964, 28, 304-310.
- HELLER, K. Ambiguity in the interview interaction. In J. M. Shlien (Ed.), *Research in psychotherapy*. Vol. 3. Washington, D. C.: American Psychological Association, 1968.
- HELLER, K., & MARLATT, G. A. Verbal conditioning, behavior therapy and behavior change: Some problems in extrapolation. In C. M. Franks (Ed.), *Behavior therapy: Appraisal and status*. New York: McGraw-Hill, 1969.
- HILGARD, E. R., & BOWER, G. H. *Theories of learning*. (3rd ed.) New York: Appleton-Century-Crofts, 1966.
- HOVLAND, C. I. Reconciling conflicting results derived from experimental and survey studies of attitude change. *American Psychologist*, 1959, 14, 8-17.
- JACOBSON, E. A. A comparison of the effects of instructions and models upon interview behavior of high-dependent and low-dependent subjects.

- Unpublished doctoral dissertation, Indiana University, 1968.
- KIESLER, D. J. Some myths of psychotherapy research and the search for a paradigm. *Psychological Bulletin*, 1966, 65, 110-136.
- MARLATT, G. A. Vicarious and direct reinforcement control of verbal behavior in an interview setting. Unpublished doctoral dissertation, Indiana University, 1968. (a)
- MARLATT, G. A. Exposure to a model and task ambiguity as determinants of verbal behavior in an interview. Paper presented at the meeting of the Western Psychological Association, San Diego, March 1968. (b)
- MARLATT, G. A., JACOBSON, E. A., JOHNSON, D. L., & MORRICE, D. J. Effects of exposure to a model receiving varied informational feedback upon consequent behavior in an interview. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, May 1966.
- MCGUIRE, W. J. Personality and susceptibility to social influence. In E. F. Borgatta & W. W. Lambert (Eds.), *Handbook of personality theory and research*. Chicago: Rand McNally, 1968.
- SARASON, I. G., & GANZER, V. J. Social influence techniques in clinical and community psychology. In C. D. Spielberger (Ed.), *Current topics in clinical and community psychology*. New York: Academic Press, in press.
- TRUAX, C. B., WARGO, D. G., CAREHUFF, R. R., KODMAN, F., JR., & MOLES, E. A. Changes in self-concepts during group psychotherapy as a function of alternate sessions and vicarious therapy pretraining in institutionalized mental patients and juvenile delinquents. *Journal of Consulting Psychology*, 1966, 30, 309-314.
- WHALEN, C. Effects of a model and instructions on group verbal behaviors. *Journal of Consulting and Clinical Psychology*, 1969, 33, 509-521.
- ZERFAS, P. G. Effects of induced expectancies and therapist activity upon patient behavior in group psychotherapy. Unpublished doctoral dissertation, Indiana University, 1965.

(Received June 6, 1969)

CONDITIONING AND TRANSFER OF AFFECTIVE SELF-REFERENCES IN A ROLE-PLAYED COUNSELING INTERVIEW¹

ROBERT J. HOFFNUNG²

University of Cincinnati

The differential effects of five forms of therapylike intervention (Um-Hmm, Echoic, Paraphrasic, combined Um-Hmm-Echoic, and combined Um-Hmm-Paraphrasic) in the conditioning and transfer of affective self-references (ASRs) were tested in a 40-minute role-played initial counseling interview. The Ss were 72 female student volunteers. Conditioning of ASRs and transfer to a TAT story telling task were found, although hypothesized relationships between discriminative cue potency of the interventions and performance were only partially supported. An explanation involving the presumed interaction of S awareness of her own self-disclosures with intervention type, degree of threat, and conditioning performance was proposed.

Recent interest in the use of verbal operant conditioning (VOC) as a research paradigm for isolating and manipulating variables thought to be important to the therapeutic interaction has led to a number of studies of the effects of interviewer interventions in relatively unstructured interview analogues (Kennedy & Zimmer, 1968; Merbaum, 1963; Merbaum & Southwell, 1965; Powell, 1968; Waskow, 1962). While emission rates of various interviewee responses have been successfully manipulated, post hoc explanations of differential intervention effectiveness have not been subjected to adequate experimental testing.

One such explanation (Merbaum, 1963) suggests that the reinforcement value of an intervention is dependent on how effective it is in providing cues which allow S to discriminate the response-reinforcement contingency. Such discriminative cues are thought to include the interruption of S's conversation, the presence of additional associative labels, and characteristics which orient S to such rewarding properties of E as attention, interest, and warmth.

¹This paper is based on a doctoral dissertation submitted to the Department of Psychology at the University of Cincinnati under the direction of Leonard D. Goodstein, whose help is gratefully acknowledged.

²Requests for reprints should be sent to Robert J. Hoffnung, Counseling Service, University of Cincinnati, Cincinnati, Ohio 45221.

Concomitant with the application of the VOC paradigm to the interview have been attempts to demonstrate transfer of conditioned changes in rates of responding to new settings. Since transfer or generalization of changes occurring within the therapeutic interview to other life situations is an important goal of most therapies, demonstration of transfer of conditioning within the VOC paradigm is highly desirable if not essential if the model is to have continued utility. With the exception of Dicken and Fordham (1967) who conditioned positive self-reference statements and statements of positive affect and reported transfer effects for selected CPI scores, there has been little evidence that conditioning achieved in an unstructured interview situation can be transferred to new settings (Lanyon, 1967; Moos, 1963; Rogers, 1960). In no case have the differential effects of interviewer intervention on both the conditioning and transfer of a response in an unstructured interview situation been reported.

The present study systematically explored the differential effects of five forms of therapylike intervention on the conditioning and transfer of affective self-reference (ASR) responses in a full length, role-played, initial counseling interview. Interventions were designed to provide differing levels of discriminative cue potency, and were, from least to most potent, as follows: Condition I: Um-Hmm; Condition II: Echoic, in which E repeated/reflected the content and mood of

S's ASR; Condition III: Paraphrasic, in which *E* restated/rephrased S's ASR response, substituting an appropriate synonym for each "feeling" word uttered, while retaining both the meaning and feeling of S's statement; Condition IV: Combined Um-Hmm-Echoic; and Condition V: Combined Um-Hmm-Paraphrasic, which represented the highest level of discriminative cue potency.

Role-playing instructions were designed to facilitate realism, and the ASR response class was selected because of the widely held view that a client's verbalizations of his personal feelings are a key component of the therapeutic relationship. Finally, the transfer task was arranged to maximize continuity in cues with the conditioning interview setting.

The following predictions were made: (a) With the exception of the Um-Hmm intervention condition (assumed to be the least potent), all interventions would produce conditioning of ASRs, and that these effects would parallel the order of predicted intervention potency. (b) Transfer would occur as a direct function of the degree of conditioning in each intervention condition.

METHOD

The 72 female University of Cincinnati introductory psychology students who served as Ss were randomly assigned to six treatment groups of 12 each, five experimental conditions, and one control condition.

Role-playing interviews were individually conducted and consisted of a 10-minute operant period during which only a small number of noncritical interventions were employed to maintain satisfactory conversational rapport followed by a 30-minute acquisition of conditioning period during which a single *E* employed one of the five interventions immediately following each ASR uttered by S. The ASR response class is described in detail by Salzinger and Pisoni (1960) and included any statement by S describing or evaluating her own state, other than intellectual or physiological, and included only those affect statements which begun with the pronoun "I" or "We."

Transfer Task

Immediately preceding and following the role-playing, S was required to tell three 2-minute TAT stories, three different cards serving as pretransfer and three as posttransfer stimuli. Cards were equated for their ability to elicit affect responses and presentations were counterbalanced. The criterion response for the transfer task included both "other" and self-

referred affective reference statements based on the assumption that affective references (ARs) to others reflect projected references to self.

To control for the influence of training and elapsed time on AR responding, independent of conditioning, Ss in Condition VI received the same pre- and post-role-playing tasks but in lieu of the conditioning interview and participated in a 40-minute "self-interview" during which they discussed 10 general topics related to their personal experiences, recording themselves in the absence of *E*.

Posttransfer Interview

Following the transfer task, each S was asked to complete a self-administered questionnaire (similar to that employed by Merbaum & Southwell, 1965) for another psychologist interested in reactions to this experiment. She was then asked to state the purpose of the interview and her evidence for this; to list all of the feeling words she could remember expressing about herself during the interview and all of those she remembered *E* expressing; to judge the pleasantness of *E*'s interventions on an 11-point Likert-type scale; and finally, to list all younger and older siblings.

RESULTS

Typed transcripts of the final 2 minutes of the first (operant), second, third, and fourth 10-minute segments of each recorded interview were coded and independently scored by *E* and two assistants. Reliability for the three scorers was .889 as estimated by the Ebel (1951) variance ratios for intraclass correlations.

Conditioning Data

An analysis of variance yielded a significant segments effect ($F = 2.28$, $df = 3/165$, $p < .05$) indicating that ASRs increased for all experimental conditions. Neither treatments nor Treatments \times Segments effects were significant. Quadratic trends in main effects for segments were significant ($p < .01$) both for overall treatments and for individual treatment groups, and from Figure 1 it can be seen that acquisition did not positively increase over all segments, but rather peaked and declined.

Due to differences in operant response level among the five treatment groups, an analysis of variance based on change scores was performed. A significant treatments effect ($F = 1.68$, $df = 4/55$, $.05 < p < .10$) indicated that with operant levels held constant, interven-

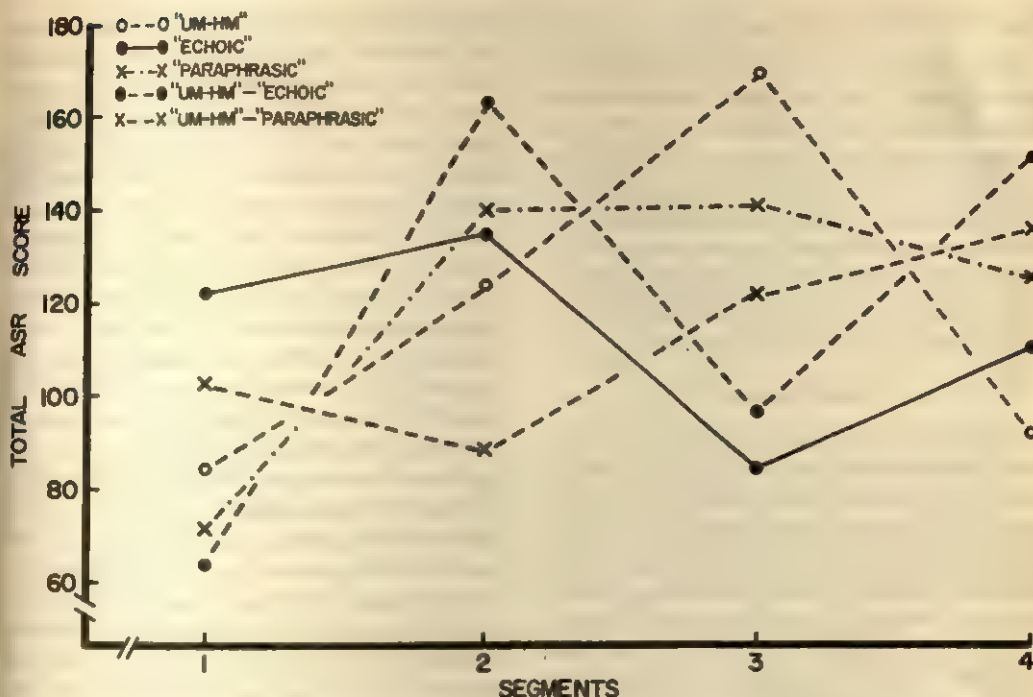


FIG. 1. Total ASR scores summed over scorers, by segments, for each of five *E* intervention conditions.

tions were to some degree differentially effective. A Newman-Keuls test of differences between treatment means revealed that the combined Um-Hmm-Echoic, the Paraphrasic, and the Um-Hmm interventions produced significantly greater conditioning than did the Echoic intervention, ($p < .01$), and that the combined Um-Hmm-Echoic and the Paraphrasic interventions were also superior to the combined Um-Hmm-Paraphrasic intervention, ($p < .05$).

Transfer Data

A transfer score was derived for each *S* by subtracting the number of affective references (ARs) given in the first set of TAT stories from the number given in the second set and transforming to eliminate negative signs. The means and standard deviations for the five experimental and one control groups are presented in Table I.

Story length and verbalization time per story remained fairly constant for pretransfer and posttransfer tasks and did not differ significantly between conditions. An analysis of variance revealed a significant treatments

effect ($F = 2.38$, $df = 5/66$, $p < .05$), and comparison of all experimental means with that of the control by Dunnett's test (Winer, 1962) showed that all treatment conditions produced greater transfer of ARs ($p < .05$) than did the control condition. The superiority of the Paraphrasic intervention over the Echoic intervention represented the only significant difference in transfer effects between treatments ($t = 4.86$, $df = 22$, $p < .05$). Although total conditioning effects (ASRs) were not significantly correlated with transfer effects (ARs), a rank-order correlation of .70 between mean conditioning and transfer scores was obtained.

TABLE I
MEANS AND STANDARD DEVIATIONS OF AFFECTIVE
REFERENCE TRANSFER SCORES

Statistic	Condition					
	I	II	III	IV	V	VI
<i>M</i>	14.83	12.75	16.50	13.50	14.25	9.75
<i>SD</i>	5.31	4.21	5.45	5.38	4.44	4.37

Postexperimental Interview

No relationships were found between birth order, family size, ratings of intervention pleasantness or conditioning or transfer performance. Only five Ss were able to verbalize the nature of the reinforcement contingency and those were distributed throughout the five intervention conditions. A gross rating of Ss' involvement in role playing indicated that approximately $\frac{3}{5}$ were highly involved, $\frac{2}{5}$ moderately, and $\frac{1}{5}$ minimally involved.

DISCUSSION

As predicted, conditioning effects occurred in all experimental conditions. While predictions of differential intervention effectiveness based on hypothesized differences in their discriminative cue potency were supported in part, the absence of more clear-cut findings suggests that an alternative explanation is needed. A number of researchers have proposed that awareness of the response-reinforcement contingency may influence VOC performance, although there is considerable controversy about whether awareness is a learned prerequisite for conditioning or merely a behavioral concomitant of, but not a necessary condition for, its occurrence (e.g., Dixon & Oakes, 1965; Spielberger & DeNike, 1966).

Although few Ss were able to report awareness during the brief postexperimental interview, the accuracy with which they recalled both their own ASRs and the interventions of *E* suggested that more were aware and that this would have been confirmed by a more detailed interview procedure. The effects of awareness on consequent performance may have been influenced by the nature of S's responses (the things she heard herself saying), the type of intervention employed (what she heard *E* saying), and such situational factors as S expectations, instructional set, and degree of situational threat. Most Ss apparently accepted as valid the role-playing instructions and appeared willing to talk about personal feelings from the outset. With the exception of the Paraphrasic intervention group, ASR responding increased for all conditions during the first conditioning segment.

Awareness would have exposed S to her own ASRs and, with the exception of Ss in

the Um-Hmm condition, confronted her with *E*'s reflections or restatements of these same self-disclosures. Awareness of such self-disclosures is often threatening particularly when negative or self-critical feelings are involved. Categorization of ASR data into positive, neutral, or negative responses revealed that negative self-disclosures predominated, the ratios being $\frac{2}{5}$, $\frac{1}{5}$, $\frac{3}{5}$, respectively. For the aware S, such feedback would probably have been threatening and led to suppression or partial extinction of ASR responding. If interventions most effective in providing cues for discriminating the response-reinforcement contingency lead to the most rapid conditioning, and if awareness and conditioning are concomitant, awareness and its consequences should have occurred most rapidly for the most potent interventions.

From Figure 1 it can be seen that Um-Hmm took the longest to produce peak conditioning while acquisition was more quickly achieved with the other more potent interventions. For the Um-Hmm, Echoic, and Um-Hmm-Echoic interventions, peak conditioning was followed by one or more decreases in response rate, suggesting that at that point, something interfered with the acquisition process. Anecdotal reports of Ss in the two Echoic conditions indicated that they found it annoying and unpleasant to hear *E* just repeat and parrot back the things they had said about themselves. This supports the notion that S's awareness of her own responses, or of *E*'s interventions was threatening. The possibility that reflections may have aversive qualities was raised by Merbaum and Southwell (1965) who found Echoic interventions to be ineffective as reinforcers. The relative stability of conditioning effects in the Paraphrasic and Um-Hmm-Paraphrasic intervention groups suggests that they were at least threatening in terms of the affects of awareness on performance. In both, the parroting quality of the Echoic interventions was mediated by a rephrasing of S's content, and in one, by the addition of Um-Hmm.

Turning to the transfer data, we find that transfer occurred for all experimental groups but not for the comparison control group, strongly supporting the second experimental

hypothesis. The finding also supports the notion that transfer is largely dependent on continuity between cues in conditioning and transfer settings. Such continuity was maximized by having the same *E* administer both the conditioning and transfer tasks, in the same room, with minimal delay between tasks. The relative insensitivity of transfer effects to the variability in conditioning performance may have been due in part to the restricted range of the transfer scores and in part to the projective nature of the TAT task which allowed *S* to attribute her feelings to characters in her stories and thereby defend against the threat of self-disclosure.

Variability in performance due to awareness and threat is characteristic of other life situations. For example, in counseling where facilitation of client expression and awareness of personal feelings is often a desired goal, the blocking and resistance frequently encountered may be analogous to the shifts in ASR responding found in the present study. Future research should investigate the proposed relationships between intervention potency, response class, awareness of threat, and conditioning in situations resembling the counseling interview.

REFERENCES

- DICKEN, C., & FORDHAM, M. Effects of reinforcement of self-references in quasi-therapeutic interviews. *Journal of Counseling Psychology*, 1967, 14, 145-152.
- DIXON, P. W., & OAKES, W. F. Effect of intertrial activity on the relationship between awareness and verbal operant conditioning. *Journal of Experimental Psychology*, 1965, 69, 152-157.
- EBEL, R. L. Estimation of the reliability of ratings. *Psychometrika*, 1951, 16, 407-423.
- KENNEDY, J. J., & ZIMMER, J. M. Reinforcing value of five stimulus conditions in a quasi-counseling situation. *Journal of Counseling Psychology*, 1968, 15, 357-362.
- LANYON, R. I. Verbal conditioning: Transfer of training in a therapy-like situation. *Journal of Abnormal Psychology*, 1967, 72, 30-34.
- MERBAUM, M. The conditioning of affective self-references by three classes of generalized reinforcers. *Journal of Personality*, 1963, 31, 179-191.
- MERBAUM, M., & SOUTHWELL, E. A. Conditioning of affective self-references as a function of the discriminative characteristics of experimenter intervention. *Journal of Abnormal Psychology*, 1965, 70, 180-187.
- MOOS, R. H. The retention and generalization of operant conditioning effects in an interview situation. *Journal of Abnormal and Social Psychology*, 1963, 66, 52-58.
- POWELL, W. J., JR. Differential effectiveness of interviewer interventions in an experimental interview. *Journal of Consulting and Clinical Psychology*, 1968, 32, 210-215.
- ROGERS, M. J. Operant conditioning in a quasi-therapy setting. *Journal of Abnormal and Social Psychology*, 1960, 60, 247-252.
- SALZINGER, K., & PISONI, S. Reinforcement of verbal affect responses of normal subjects during the interview. *Journal of Abnormal and Social Psychology*, 1960, 60, 127-130.
- SPIELBERGER, C. D., & DENIKE, L. D. Descriptive behaviorism versus cognitive theory in verbal operant conditioning. *Psychological Review*, 1966, 73, 306-326.
- WASKOW, I. E. Reinforcement in a therapy-like situation through selective responding to feelings or content. *Journal of Consulting Psychology*, 1962, 26, 11-19.
- WINER, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

(Received August 21, 1968)

ADJUSTMENT, MENTAL HEALTH OPINIONS, AND PROFICIENCY OF CHILD CARE PERSONNEL¹

JOHN W. GIEBINK²

AND

DONALD O. STOVER

University of Wisconsin

Wisconsin Children's Treatment Center

MMPI validity and *K*-corrected diagnostic scale scores, responses to a mental health opinion questionnaire, biographical data, and job performance ratings of 52 child care personnel were intercorrelated. There were significant ($p < .05$) relationships between 11 MMPI scales and measures of job proficiency, but only 4 of the 40 mental health opinion items were related to proficiency. A principal components analysis of the entire correlation matrix yielded two major factors. Items related to adjustment and job proficiency were strongly loaded on the first factor. The second factor reflected the biographical variables. The remaining factors accounted for small portions of the total variance and seemed to be relatively specific in nature. Since mental health opinions in contrast to personal adjustment were seemingly unrelated to job proficiency, it was suggested that their "mental illness" determined content was irrelevant to performance on a job requiring a "learning" orientation.

An assumption often made but seldom tested is that personal adjustment affects the performance of persons caring for emotionally disturbed children and adults. Both adjustment and mental health attitudes are frequently considered important criteria in selecting staff to care for emotionally disturbed persons. These two qualities would seem to be especially related to the job proficiency of child care workers in an intensive residential treatment center, for their work requires the ability to enter into close therapeutic relationships with children whose objectionable behavior has resulted in separation from home and community.

Although studies suggest that certain personality and attitudinal attributes may be altered by contact with persons confined in mental hospitals (e.g., Holzberg & Gewirtz, 1963; Holzberg, Gewirtz, & Ebner, 1964), evidence concerning the relationships between existent personality and attitudinal characteristics and the behavior of care personnel is sparse. After comparing psychiatric aides

with normative groups on the Edwards Personal Preference Schedule and the Leary Interpersonal Checklist, Lawton (1964) attempted to relate these measures to their job performance. Although performance ratings were "unstable" and interjudge agreement was low, he described the better aides as less authoritarian, more naively benevolent, and as seeing themselves as less discrepant from the patients than the poor aides. Less than the chance number of significant correlations were obtained in a study relating opinions about mental illness, interpersonal values, and some biographical variables to factor scores and total scores of performance ratings of psychiatric aides (Tollefson, 1964). A factor analysis of all the variables in this study yielded no factors with strong loadings from both predictor and criterion measures. Pryer (1967) found no relationship between three objective personality measures on psychiatric aides and ratings of their job performance.

There are some possible explanations for the lack of empirical support for the assumption that there would be a relationship between personality and/or attitudes and job performance in caring for disturbed persons. It is quite likely that persons with markedly deviant personality attributes as well as those who hold unusual opinions about mental health will be screened out prior to employment, thus restricting the range of scores on

¹ Partial support for this study was provided by a grant from the University of Wisconsin Graduate School. Julianne Murphy, Director of Nursing at the Children's Treatment Center, facilitated this research and did much of the performance rating. Stan Rubinstein assisted in the data analysis.

² Requests for reprints should be sent to John W. Giebink, Psycho-Educational Clinic, University of Wisconsin, Madison, Wisconsin 53706.

these attributes and consequently decreasing the chances of finding significant relationships. Instability of the personality and attitudinal measurements and the lack of reliability of ratings of performance would also limit the correlations obtained.

The rather meagre results reported previously along with the methodological difficulties noted above suggested the use of procedures that would enhance the chances of finding relationships among adjustment, opinion, and performance measures. This could be done by obtaining reliable performance ratings and using a correlational approach to maximize the opportunity of finding a statistical relationship among MMPI scale scores, responses to a mental health opinion questionnaire, certain biographic variables, and performance ratings.

METHOD

The Minnesota Multiphasic Personality Inventory (MMPI) and a mental health opinion questionnaire

(Nunnally, 1961) were routinely administered to all nurses and child care workers within 2 weeks of their employment at the Wisconsin Children's Treatment Center (a facility for research and treatment for severely disturbed children). Although the mental health opinion questionnaire consisted of 60 items, only those 40 items that had been factored by Nunnally were used in the present study. Each of these items was rated on a 7-point "agree-disagree" scale. The Ss' performance on these measures was not available to the staff at the Center, and consequently could not enter into decisions about selection, evaluation, or retention of employees.

Three evaluations of job performance were available for the 52 Ss in this study. One of these, the Specific Proficiency Rating (SPR) consisted of 98 true-false items. After 6 months' employment, the SPR was completed on each staff member by his nursing supervisor. The Hoyt (1941) internal consistency estimate of reliability for this scale was .87. The other two performance measures consisted of global ratings made on 5-point scales by the director of nursing. The Retention Rating was a judgment of the amount of effort that either was expended or would potentially be expended to retain the employee. The Global Proficiency Rating was a retrospective judgment of the overall effectiveness

TABLE 1
CORRELATIONS BETWEEN PREDICTOR AND CRITERION VARIABLES

Variables	Specific proficiency rating	Retention rating	Global proficiency rating
MMPI			
L	-.789	-.708	
K	?		-.426
F	-.684	-.722	
Hs	-.795	-.644	
D		.408	
Hy	-.707	-.602	
Pd	-.557	-.441	
Pa	-.805	-.690	
Pt	-.748	-.583	
Sc	-.472	-.326	
Ma	-.792	-.714	
Si	-.690	-.630	
Biographic			
Child care worker—nurse			.329
Mental health opinions			
Women have no more emotional problems than men do.			-.383
X rays of the head will not tell whether a person is likely to become insane.		-.302	
If a person concentrates on happy memories, he will not be bothered by unpleasant things in the present.		-.295	
The good psychiatrist acts like a father to his patients.	-.400		
Almost any disease that attacks the nervous system is likely to bring on insanity.	-.301		

Note.—Only those correlations which were significant ($p < .05$, $r = .273$, $df = 50$) are reported.

TABLE 2
ITEM LOADINGS FOR FACTORS I AND II
FOLLOWING VARIMAX ROTATION

Item	Factor loading
Factor I	
MMPI	
Pa	-.961
Hs	-.949
Pt	-.938
L	-.926
Hy	-.920
Ma	-.918
F	-.897
Specific Proficiency Rating	.842
MMPI	
Si	-.784
Sc	-.748
Global Retention Rating	.741
MMPI	
PD	-.703
Factor II	
Male-Female	.896
Position (child care worker, nurse)	.886
Degree (BA or RN)	-.839
Education beyond high school	.810
MMPI	
MF	-.762
Age	-.477
Mental health opinion items	
The mentally ill pay little attention to their personal appearance.	-.347
Mental health is one of the most important national problems.	-.347
You can tell a person who is mentally ill from his appearance.	-.346
Helping the mentally ill person with his financial and social problems often improves his condition.	-.317

of the employee at the end of his first 6 months' employment.

The biographic variables of sex, age, position (nurse or child care worker), and education, along with the three proficiency ratings, three MMPI validity scale scores and the 10 *K*-corrected diagnostic scale scores, and 40 mental health opinion items were included in the total analysis.

RESULTS AND DISCUSSION

An initial analysis consisted of the inter-correlations of all variables. The relationships among the three criterion variables of job proficiency were examined first. The Retention Rating correlated .70 with the Specific

Performance Rating, and .54 with the Global Performance Rating. The Global Performance Rating correlated .05 with the Specific Performance Rating. Thus, the Global Performance Rating appeared to be the least related to the other measures and consequently least satisfactory.

The correlation matrix was next inspected to determine the strengths of the relationships of biographic, MMPI, and mental health opinion scores to the job proficiency measures. Those relationships which resulted in statistically significant ($p < .05$) correlations are presented in Table 1. Most of the MMPI scales are clearly related to the Specific Performance and Retention Ratings, while only a few of the mental health opinion items were so related. The strength and direction of the correlations between the personality variables and job performance indicates that adjustment is related to job proficiency. Few significant correlations between the Global Proficiency Rating and other variables were obtained.

In order to further clarify the relationships among the variables, the correlation matrix was subjected to a principal components analysis. Nineteen components with eigenvalues greater than one were subjected to orthogonal rotation according to Kaiser's (1958) Varimax procedure. These 19 factors accounted for 83.23% of the total variance. The first factor accounted for 15.09% of the total variance, the second accounted for 10.18%, and the third for 5.09%. The remaining 16 factors accounted for less than 5% each. Items loading .30 or higher on the first two factors are presented in Table 2.

The first factor is composed of the major MMPI scales, the Specific Proficiency Rating, and the Global Retention Rating. The magnitude of the loadings of the proficiency scales and the MMPI scores is impressive. Elevations on any of these MMPI scales would indicate psychological disturbance. Consequently, their combination with the proficiency scores on the first factor again supports the hypothesized relationship between general psychological adjustment and competence in caring for severely disturbed children.

The second factor appears to be determined by the *S* sampling. The item loadings indicate that *Ss* were comprised of nurses and child

care workers, and reflects the fact that child care workers were males working toward college degrees, while the female nurses already had their degrees and were somewhat older. Four mental health opinion items also appeared on this factor.

Each of the rest of the several factors accounted for a small percentage of the total variance, and for the most part they were comprised of mental health opinion items. Only a few items loaded very heavily on each of these factors and they consequently seemed to be relatively specific in nature. This multiplicity of factors is consistent with the fact that the mental health opinion items were contained in 10 factors in Nunnally's original work. The Specific Proficiency Rating did not load on any of these remaining 17 factors, while the Retention Rating and Global Proficiency Rating each loaded slightly on a couple of them.

In contrast to the results of previous studies, the present findings support the assumption of a relationship between general psychological adjustment and child care proficiency in a setting demanding the formation of meaningful relationships with severely disturbed children. Demonstration of this relationship may be a consequence of using stable measures of job performance and personality. Admittedly the use of a large number of variables would enhance the chances of finding significant relationships, and the factor analytic procedure would make maximum use of minimal individual differences. Nevertheless, the clarity of the relationship is impressive, for the *S* sample represented a restricted adjustment range because of preemployment screening.

Although any correlational study is limited by the items that have been selected for use, the resulting correlations, and in this case the factor analytic solution, is informative of the kinds of items that are most associated with one another. Thus, it is important to note those items that did not seem to be related to job proficiency as well as those which were found to be related. It does seem that the mental health opinions held by these persons at the time of their employment were relatively unrelated to their ultimate job proficiency.

The use of individual item scores which would be expected to have lower reliability than scores based on several items may have contributed to the low correlations. However, another possible explanation for this lack of a relationship derives from the content of the mental health opinion questionnaire itself. An examination of the questionnaire reveals that it is more concerned with "mental illness" than with mental health; many of the items deal with the etiology and treatment of "mental illness." As has been pointed out frequently (e.g., Sarbin, 1967), this model may not be pertinent. Attitudes based on the medical model could be irrelevant to the kinds of behavior which are important for the psychological treatment of disturbed children. This point may have been underscored in the present study since the nurses and child care workers were rated for their performance in a facility which is little influenced by the mental illness concept, where personnel are highly aware of "behavior," and where learning principles are used to guide treatment procedures.

REFERENCES

- HOLZBERG, J. D., & GEWIRTZ, H. A method of altering attitudes toward mental illness. *Psychiatric Quarterly Supplement*, 1963, 37, 56-61.
- HOLZBERG, J. D., GEWIRTZ, H., & EBNER, E. Changes in moral judgment and self-acceptance in college students as a function of companionship with hospitalized mental patients. *Journal of Consulting Psychology*, 1964, 28, 299-303.
- HOYT, C. J. Test reliability estimated by analysis of variance. *Psychometrika*, 1941, 6, 153-160.
- KAISER, H. F. Varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, 23, 187-200.
- LAWTON, M. P. Studies on the psychiatric aide. *Mental Hospitals*, 1964, 15, 512-515.
- NUNNALLY, J. C. *Popular conceptions of mental health*. New York: Holt, Rinehart & Winston, 1961.
- PRYER, R. S. Basic issues involved in attendant selection, training, and evaluation. *Public Personnel Review*, 1967, 28, 46-50.
- SARBIN, T. R. On the futility of the proposition that some people be labeled "Mentally Ill." *Journal of Consulting Psychology*, 1967, 31, 447-453.
- TOLLEFSON, T. I. The relation between attitudes, personal values, biographical data and proficiency ratings obtained upon psychiatric aides. *Dissertation Abstracts*, 1964, 24, 5552.

(Received August 23, 1968)

THERAPIST DIRECTIVENESS, CLIENT DOMINANCE, AND THERAPY RESISTANCE

ROBERT W. HAGEBAK

AND

GEORGE V. C. PARKER¹

Timberlawn Psychiatric Center

University of Texas at Austin

This therapy analogue study dealt with the interaction of client personality (dominance—Dom) and therapist directive versus nondirective techniques in producing client resistance. Forty High-Dom and 40 Low-Dom undergraduate male Ss, while role-playing client problems, were required to make choices between 32 pairs of high- and low-resistant client responses following therapist statements. Results showed that nondirective techniques elicited most client resistance, Low-Dom Ss tended to give more resistant responses, and both effects were most pronounced when Ss assumed hostile roles. The S anxiety was also found indirectly to be related to resistance.

Familiar to virtually all therapists is patient behavior during the therapy interview which appears to be uncooperative and even negativistic with regard to the goals of therapy. Generally conceived, this is one aspect of the classical problem of "resistance," a concept born within the framework of traditional psychoanalytic theory (cf. Fenichel, 1945, p. 26) but which has come to possess a variety of etiological connotations, including clients' fear of social disapproval and therapists' resentment about patients' noncooperation. Regardless of one's theoretical allegiance, the significance of the resistance phenomenon in therapy is hard to overestimate. Indeed, Schwebal (1960) has remarked that "The various investigations of the correlates of success or failure, of continuing or terminating in counseling and therapy, may well be separating the high and low resisters [p. 482]." The position taken here is that understanding of resistance in therapy may be advanced by conceptualizing this behavior as a function of client personality variables, and of differential client responses to various therapeutic techniques.

Classified broadly, empirical studies of resistance have come from two directions which could probably benefit from integration: (a) those that relate resistive client behavior to client and therapist characteristics defined psychometrically (e.g., Heilbrun, 1961a, 1961b; Taulbee, 1958) and (b) those which

relate resistive client behavior to directive and nondirective therapist interview behavior (e.g., Bergman, 1951; Carnes & Robinson, 1948; Frank, 1964; Frank & Sweetland, 1962; Truax, 1961, 1963). While these studies are important and provocative, they do not permit evaluation of the interactions of the combined roles of client personality and therapist interview behavior in producing client resistive behavior. Based on findings of these earlier investigations, two major hypotheses were generated: Hypothesis I: High-dominance Ss would offer more resistance in a therapy analogue situation than would low-dominance Ss. Hypothesis II: Directive therapist statements would elicit more resistant client responses in a therapy analogue situation than would nondirective therapist statements.

METHOD

Subjects

High-dominance Ss. Forty Ss were selected from a larger pool of approximately 240 male undergraduates at the University of Texas, who were given the Adjective Check List (ACL; Gough & Heilbrun, 1965). Those with highest scores (*T* of 60 or above) on the Dominance scale (Dom) of the ACL represented the High-Dom group (High Dom).

Low-dominance Ss. Forty Ss were similarly selected from the same pool on the basis of having the lowest Dom scores (*T* of 44 or below) on the ACL (Low Dom).

Test Stimuli

Typescripts. During the course of the experiment, Ss were asked to assume the roles of eight hypothetical clients (two each with academic, interpersonal anxiety, homosexual, and hostile impulse control problems) who had sought professional help.

¹Requests for reprints should be sent to George Parker, Department of Psychology, University of Texas, Austin, Texas 78712.

All Ss were instructed to respond to the therapist statements in a manner consistent with what they believed their own behavior would be in each situation and were informed that there were no "right" or "wrong" answers.

A short description of the problem leading the client to seek professional help was followed by four typical therapist statements, each of which was followed by a choice of two client responses, one resistant, the other a nonresistant, cooperative response. The Ss responded to directive therapist statements in one of the two cases in each problem area, and to nondirective statements in the other case of the same problem area. Orders of presentation of therapist statements and the problem areas were counterbalanced to minimize sequence effects.

Similarly, the order of the resistant-nonresistant client alternative responses was alternated in each case.

Therapist Statements

As in other studies of therapist directiveness-non-directiveness (Bohn, 1965; Carnes & Robinson, 1948; Finesinger, 1948; Frank, 1964; Parker, 1967; Parsons & Parker, 1968), directive therapist statements were defined as those which would tend clearly to lead, direct, or control the verbal activity during the therapy session, and included direct questions and forcing the topic directly. Nondirective therapist statements were defined as those which would tend to give responsibility of decision for choice of area and direction of verbal activity largely to the client, as well as those responses which reflect or clarify the client's affect. Nondirective leads, clarification of feeling, and restatement of content were the three types of nondirective therapist statements used. The classification of these statements in the analogue material was assessed by asking three judges (experienced therapists) to categorize each therapist statement. Interjudge agreement with regard to the major directive-nondirective categories was high (94%).

Client Responses

Resistant client responses were defined as those indicating overt or covert opposition to the therapist and the therapeutic process and/or an unwillingness to explore problems, to establish rapport, or to make collaborative efforts. Here also, a similar group of three judges was asked to indicate which alternative in each pair of client responses was resistant and which was nonresistant. One hundred percent interjudge agreement was achieved.

Ability of Ss to Assume Roles

Debriefing indicated that Ss generally had little difficulty assuming the client roles. However, in order to obtain a more systematic assessment of Ss' ability to role-play, a postexperiment questionnaire was given. This focused on Ss' feelings of nervousness during the experimental task, and was evaluated in relation to S Dom level, as described subsequently.

TABLE 1

ANALYSIS OF VARIANCE OF RESISTANT RESPONSES AS A FUNCTION OF "CLIENT" DOMINANCE, THERAPIST DIRECTIVENESS, AND TYPE OF CLIENT PROBLEM

Source	df	F	p
Between	79		
Dominance (A)	1	3.15	.08
Error	78		
Within	560		
Directiveness (B)	1	7.28	.008**
A × B	1	2.25	.13
Error	78		
Client Problems (C)	3	1.21	.30
A × C	3	.36	.79
Error	234		
B × C	3	3.24	.02*
A × B × C	3	.24	.87
Error	234		
Total	639		

* $p < .05$.

** $p < .01$.

RESULTS

Hypotheses I and II were tested with a Type VI analysis of variance (Lindquist, 1953). The between-groups variable was dominance while directiveness and type of client problem were the two within-groups factors. Results of the analysis are summarized in Table 1.

Surprisingly, the overall results were consistent with the reverse of both hypotheses. As shown in Table 1, there was a nonsignificant tendency ($p < .08$) for Low-Dom Ss to give more resistant responses than High-Dom Ss. In relation to Hypothesis II, the results were highly significant ($p < .008$) but directly contrary to those predicted, with more resistant responses following nondirective therapist statements than after directive statements.

A third finding which was not predicted was the significant interaction ($p < .02$) between type of client problem and therapist directiveness. This interaction is shown in Figure 1, from which it can be seen that nondirective therapist statements were followed by more resistant responses than were directive statements in all client problem areas except academic.

For academic problems, while not statistically significant ($t = 1.33$, $p > .05$), this re-

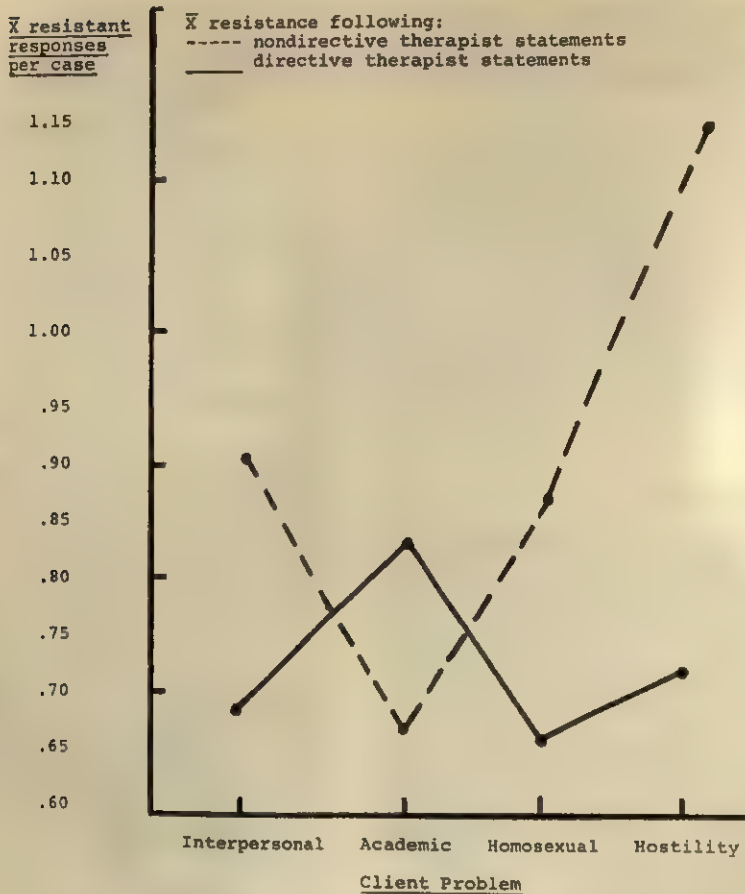


FIG. 1. Interaction of therapist directiveness with problem type.

relationship was reversed. The mean numbers of resistant client responses per case following directive therapist statements were compared with those for nondirective statements for each of the four problem areas. The difference was significant only for hostility control problems ($t = 5.85, p < .001$). Following this finding, an independent sample of 20 undergraduate Ss was asked to rate on a 9-point scale the de-

gree of each problem's seriousness (1 = mildly serious; 5 = quite serious; 9 = extremely serious). Seriousness was defined as the extent to which the problem would (a) cause the client to experience interpersonal difficulties, (b) reduce his effectiveness in reaching personal goals, (c) interfere with his achieving a reasonable degree of personal satisfaction, and (d) create the need for professional therapeutic intervention. The mean ratings obtained were academic, 2.07; interpersonal, 5.88; homosexual, 6.65; hostility, 8.14. By t test comparisons, all differences in average ratings of seriousness were significant. ($p < .05$) except between interpersonal and homosexuality problems. The results of this analysis suggest that resistance, particularly to nondirective therapy techniques, is related to the perceived severity of the problem.

TABLE 2
NERVOUSNESS PLAYING CLIENT ROLES:
HIGH- AND LOW-DOM AVERAGES

Group	Not bothered	Felt uneasy
High Dom	30.1	9.9
Low Dom	23.6	16.4

Note.— $\chi^2 = 6.32, p < .02$.

Results of the postexperiment questionnaire summarized in Table 2 indicated that, averaging over all client roles, significantly more Low-Dom Ss than High-Dom Ss reported feeling uneasy while performing the task. As described below, it seems clear that this relationship may have considerable bearing on explanation of the Dom-resistant findings being generally the reverse of prediction.

DISCUSSION

While the results were not consistent with the specific predictions, the more general hypothesis that resistance may be a function of therapist technique and of client personality characteristics has received support. The main effect showing significantly fewer resistant responses to directive than to nondirective therapist statements must of course be qualified in the light of a significant interaction between directiveness and type of client problem. This interaction suggests that resistance may be reduced if therapist technique is varied according to the type of client problem. Closer examination of the analogue data indicated that nondirective statements elicited significantly more resistance than directive statements with the "most severe" (hostility control) problems and the least resistance (less than directive—a reversal of the main effect) with the least severe (academic) problems.

Therefore, while the main effect would suggest that these male Ss, most of whom had not had counseling or therapy, would favor a more directive, controlling therapist to one who places the responsibility for direction of the interview onto the client, this may be least likely to be true with milder complaints (i.e., academic or study problems). Considering the similarity of the present results to Klein's (1967) finding that directive advice-giving therapy was strongly preferred over nondirective, reflective therapy by naive adult Ss in an experimental situation, another major implication is that in an ambiguous situation such as an initial therapy interview, it is easier, less demanding, and perhaps less threatening to respond to direct questions and themes chosen by the therapist than to nondirective statements.

Other possible explanations of these results bear on problems related to the analogue situ-

ation itself. The Ss here were not actually seeking help for the problems they role-played. With the single exception of academic problems, the problem areas here were probably not personally significant to many of the Ss. Consequently, it could be argued that in the analogue situation, Ss were not only responding to therapists' statements but were also attempting to play a largely foreign role. Of course, in some respects, this is not unlike the situation in which an actual client finds himself when he comes for an initial interview. Any directive therapist remark that would help *S* in his task by specifying what he should talk about (as opposed to nondirective cues that would leave *S* shouldering responsibility largely on his own) would doubtless tend to be followed by "cooperative" client statements. It is not altogether clear that this kind of confounding has been successfully dealt with in most analogue studies. It is clear, however, that it is difficult to account precisely for the contribution of these factors when generalizing from the analogue to actual clinical practice.

The finding of a tendency toward a negative relationship between *S* Dom and resistance, although in reverse of prediction, is consistent with the psychoanalytic notion that greater resistance is associated with greater anxiety. The postexperiment questionnaire data in this study indicated that Low-Dom Ss felt more uneasy playing these client roles than did the High-Dom Ss and significant negative correlations have been reported between the Dom scale and the MMPI *Pt* scale as well as the Welsh Anxiety scale (Gough & Heilbrun, 1965). Perhaps the basic implication of this finding is that there is a hierarchy of factors affecting client responses in the therapy situation and these may be considered in terms of their prepotency. The evidence here suggests that client anxiety is a more powerful influence than client Dom in accounting for client behavior, at least early in the therapy relationship.

REFERENCES

- BERGMAN, D. Counseling method and client response. *Journal of Consulting Psychology*, 1951, 16, 215-244.
- BOHN, M. J., JR. Counselor behavior as a function of counselor dominance, counselor experience, and

- client type. *Journal of Counseling Psychology*, 1965, 12, 346-352.
- CARNES, H. F., & ROBINSON, F. P. The role of the client talk in the counseling interview. *Educational and Psychological Measurement*, 1948, 8, 635-644.
- FENICHEL, O. *Psychoanalytic theory of neurosis*. New York: Norton, 1945.
- FINESINGER, J. H. Psychiatric interviewing: I. Some principles and procedures in insight therapy. *American Journal of Psychiatry*, 1948, 105, 187-195.
- FRANK, G. H. The effect of directive and nondirective statements by therapists on the context of patient verbalizations. *Journal of General Psychology*, 1964, 71, 323-328.
- FRANK, G. H., & SWEETLAND, A. A study of the process of psychotherapy: The verbal interaction. *Journal of Consulting Psychology*, 1962, 26, 135-138.
- GOUGH, H. G., & HEILBRUN, A. B. *The Adjective Check List manual*. Palo Alto: Consulting Psychologists Press, 1965.
- HEILBRUN, A. B. Male and female personality correlates of early termination in counseling. *Journal of Counseling Psychology*, 1961, 8, 31-36. (a)
- HEILBRUN, A. B. Client personality patterns, counselor dominance and duration of counseling. *Psychological Reports*, 1961, 9, 15-25. (b)
- KLEIN, D. Parental preferences for counseling approaches as a function of social class. Unpublished doctoral dissertation. University of Texas, 1967.
- LINDQUIST, D. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.
- PARKER, G. V. C. Some concomitants of therapist dominance in the psychotherapy interview. *Journal of Consulting Psychology*, 1967, 31, 313-318.
- PARSONS, L., & PARKER, G. V. C. Personal attitudes, clinical appraisals, and verbal behavior of trained and untrained therapists. *Journal of Consulting and Clinical Psychology*, 1968, 32, 64-71.
- SCHWEBEL, M. Counselor intervention in resistance and ignorance. *Personnel and Guidance Journal*, 1960, 38, 480-485.
- TAULBEE, E. S. Relationship between certain personality variables and continuation in psychotherapy. *Journal of Consulting Psychology*, 1958, 22, 83-89.
- TRUAX, C. B. The process of group psychotherapy. *Psychological Monographs*, 1961, 75(7, Whole No. 511).
- TRUAX, C. B. Effective ingredients in psychotherapy: An approach to unravelling the patient-therapist interaction. *Journal of Consulting Psychology*, 1963, 10, 256-263.

(Received August 27, 1968)

TEMPORAL CORRELATES OF THE ACHIEVEMENT VALUE AND MANIFEST ANXIETY¹

THOMAS J. COTTLE²

Department of Social Relations, Harvard University

A study of the relationship between achievement values, manifest anxiety, and orientations to time was reported. Controlling for the effect of intelligence, it was found in a sample of respondents, ages 18-23, that whereas achievement is related to accepting the possibilities of the future and to perceiving an "integration" between the past, present, and future, anxiety is associated with apprehensions about the future, and functions in decreasing the sense of relatedness between time zones. In addition, anxiety is associated with fantasies of temporal recovery and preknowledge. It was suggested that the achievement value enhances a realistic acceptance of time's unalterable properties, while anxiety encourages fantasy resolutions of realistic temporal problems.

From the point of view of time, it is the concept of the future that distinguishes most clearly the achieving person from the anxious one. Whereas the achiever involves himself in mastering destiny (Strodtbeck, 1958), the anxious person seems overly apprehensive or indeed dreading the future's arrival (Persky, Grinker, Mirsky, & Gamm, 1950). Moreover, whereas achievement values reinforce the urge to keep moving forward in time (McClelland, 1961), anxiety increases one's appetite for the safety of the familiar and actually serves as a danger of impending events (Freud, 1936).

Several psychologists have demonstrated systematic variations in time perception among achieving and anxious persons (Epley & Ricks, 1963; McClelland, 1961; Wallace & Rabin, 1960; Wohlford, 1964). McClelland (1961), for example, reported that high achievers used more future tense sentence construction than low achievers of a comparable sample. Anxiety, on the other hand, has been shown to deflect individuals away from the future and increase sentiments that success in personal action has a low probability of occurrence. Wohlford (1964) observed:

The farther into his future a high anxiety individual roams on a pre-conscious level, the more he is frightened by what he sees there, and this causes him to draw back more into the present and past in his conscious thought [p. 82].

Wohlford's words are reminiscent of Freud's (1936) definition of anxiety as "the reaction to a situation of danger; [usually] circumvented by the ego's doing something to avoid the situation or retreat from it [p. 65].

Accordingly, it is not surprising that anxious persons would attribute a heightened degree of potency to the future (MacKinnon, 1944). In so doing, they may unwittingly be confessing an inability to confront the unknown or simply cope with their own apprehensions. Achievers on the other hand, seem to be governed by a sense of personal responsibility (Strodtbeck, 1958) and even vital self-preoccupation (Henry, 1949). Thus, if anxiety cause an exaggeration of the future's dangers, achievement values reinforce an exaggerated sense of man's ability to control these dangers.

One way to understand the variations in time perceptions of anxious and achieving persons is to examine the extent to which individuals relate the past, present, and future. There is, for example, a constant forward thrust in the achiever's present which results perhaps from his sense of the unattained or his frequent feelings of unfulfillment (McClelland, 1961). It is not surprising that his involvement with chronological time should be one of such eager anticipation that he sets his watch a few minutes fast (Cortes,

¹ The author wishes to thank Fred L. Strodtbeck, Jacob W. Getzels, Donald N. Levine, and John M. Butler for their comments on earlier drafts of this paper. Similar thanks are due Joseph Pleck, William Morgan, Carl N. Edwards, Linda Janus, Cynthia Sparrow, and Susan Giese.

² Requests for reprints should be sent to Thomas J. Cottle, Department of Social Relations, Harvard University, Cambridge, Massachusetts 02138.

1960, cited in McClelland, 1961). Through active planning, achievers unite the present and future. As McClelland (1951) wrote,

It is as if the [achievement] need has served to relate present achievement experiences to future ones, to promote understanding of the present in terms of a wider context. Motives seem to tie the present to the future, the specific to the general and long run [p. 486].

In contrast, anxious persons appear less able to establish this union of time zones. In many of their perceptions, they reflect a temporal as well as personal disunity (May, 1950) or what some have called ego disintegration. For them, the temporal horizon is muddled, futures are not adequately differentiated from presents (Melges & Fourgerousse, 1966), and the fantasy of regaining the past is common.

A connection between time zones, and particularly between the present and future, suggests a sense of competence in one's ability to shape the future through personal effort (Muller, 1967). If this is true, achievers should portray the connection more often than anxious persons. Equally important, confidence in affecting outcomes, being the agent of causation (Bakan, 1966) or otherwise influencing one's environment (White, 1959), should lessen the need for fantasy solutions to temporal problems. The past now may be relinquished and the future left unknown until that proper moment when the feeling of efficacy described by William James (1912, cited in Muller, 1967, p. 9) comes to be experienced:

Sustaining, persevering, striving, paying with effort as we go, hanging on and finally achieving our intention—this is pure action, this is effectuation in the only shape in which, by a pure-experience philosophy, the whereabouts of it anywhere can be discussed . . . here is causality at work.

The present study examines relationships between achievement values, manifest anxiety, and six areas of time perception. In so doing, 11 major hypotheses are generated.

The first two variables explore an experimental time orientation and establish basic temporal definitions of the present. It is hypothesized that the achiever's forward look will display itself (*a*) in an emphasis on expectation, rather than on reflection, and (*b*) in a demonstration that the future border

of the present, essentially the time between what an individual defines as now and the commencement of his future, will seem more important, hence, longer than the past border of the present, essentially the time between now and the conclusion of one's past. On the other hand, anxiety, which has been defined in terms of fear of the unknown (Freud, 1936) (*c*) should reduce the significance of expectations but (*d*) increase the degree to which past experiences become important. This past orientation should reflect in a demonstration that the present's past border, that is, the time of completed events still within the present, seems more salient, hence longer.

The notions governing these hypothesized associations, therefore, are, first, that salience and extension are equated, and, second, that the salience of time zones, as in past and future orientations, reflects in the extensions of the present's boundaries corresponding to the orientation.

Because of their belief in personal control or the sense that they will affect the future, as they have the present, achievers (*e*) should demonstrate temporal relatedness or the sense of a connection between the past, present, and future. In contrast, as a way of disentangling the confusion of time zones (Melges & Fourgerousse, 1966), anxious persons (*f*) should demonstrate temporal unrelatedness, which symbolizes, moreover, their lack of a sense of personal efficacy and control (Cottle, 1967; White, 1959).

Anxiety about the pace of life and apprehension about the future (*g*) should increase as a function of manifest anxiety but (*h*) decrease as a function of valuing achievement. This last hypothesis is advanced with the rationale that the constant establishment of new aspirations characteristic of achievers makes their future seem less ambiguous and foreboding. Furthermore, because of achievers' preference for moderate risk-taking situations (McClelland, 1961), achievers should manifest a belief that sufficient time exists to complete present activities and conceive future plans. Thus, future events should appear more manageable to achievers.

Anxious concern with the future (*i*) should cause an exaggeration of the sense of its potency (MacKinnon, 1944) and, in an almost

wishful manner, its goodness as well. But the responsibility for personal management of the future, so paramount in the achiever's outlook, (j) should lessen the potency and evaluation ratings of the future. Once again, preoccupation with an unmanageable stimulus, in this case the future, produces heightened estimations of its power and goodness.

Finally, anxiety (k) should encourage retreats to fantasy explorations of the past and future. Rationale for this hypothesis comes, first, from Freud's (1936) observation that anxiety evokes temporal retrieval fantasies which permit one to rearrange prior action or redo or undo unsatisfactory experiences. Second, in light of the apprehension generated by future unknowns, preknowledge fantasies should seem especially exciting as they offer a rather protected way of dealing with these unknowns and uncertainties.

Predicting relationships between achievement values and temporal fantasies is more difficult. Central to achievement is its minimization of prior involvements, particularly with family members (Strodtbeck, 1958). While this feature may reflect in the relatedness variable, that is, in the disconnection of the past from the present and future, it is not clear whether commitment to achievement permits the luxury of temporal retrieval through fantasy.

Fantasies of preknowledge also may cause achievers conflict as their many ideals and heightened investment in the future may make them more eager to learn how the future works out. Yet, in a more realistic framework, the mystery produced by unknowns may provide the very energy igniting the achievement drive in the first place. Thus, while anxiety should support fantasy resolutions of problems, reactions of achievers to temporal fantasies are not easily predicted.

A serious question arising in the present research was to what extent does intelligence influence perceptions of time? Although the data are not always consistent, intelligence has been shown to increase future extensions on various time measures and Thematic Apperception Test responses (Epley & Ricks, 1963; Murray, 1959; Wohlford, 1964). Furthermore, studies demonstrating interrelationships between intelligence, achievement, and

anxiety (Lavin, 1965) indicate that intelligence should be examined for the effect it might have on the two independent variables.

METHOD

Subjects

The Ss were 332 men and 98 women from a mid-west Naval Training Station. All were students in a voluntary medical corpsman program consisting of elementary medical education and patient care.

Given their parents' occupations and completed education levels, Ss were judged to be from the lower middle class. All had finished high school but few had been to college. Most came from moderate sized cities, rarely from rural areas. Although ranging in age from 18-23, less than 4% were married, and in the main, all were Protestant or Catholic. Finally, based on tests to be described, these students were found to be significantly more intelligent than their regular Navy colleagues.

One advantage of using this rather homogeneous population is that all students were engaged in practically the identical activities during the time of testing. Thus, if the program or Navy environment affected temporal perceptions, the effects should have been similar on all 430 students.

Instruments

Achievement values. Achievement values (Vach) were measured with an 8-item inventory (Strodtbeck, 1958), derived from the theory that achievement implies individual and autonomous action, personal control of destiny, and minimization of dependencies on parents.

Items are presented as statements with which respondents are either to agree or disagree. Examples of items include, "Planning only makes a person unhappy since your plans hardly ever work out anyway"; "When a man is born the success he's going to have is already in the cards, so he might as well accept it and not fight against it"; "Nowadays, with world conditions they way they are, the wise person lives for today and lets tomorrow take care of itself." The maximum Vach score is eight.

Anxiety. Anxiety was measured with the Taylor Manifest Anxiety (MA) scale (Taylor, 1953, 1956). A shortened form consisting of 21 items was used as it was found in a matched sample to correlate highly with the complete MA scale (for males, $r=.87$; for females, $r=.85$).^a The MA scale instructs Ss to agree or disagree with statements regarding self-perceptions of anxiety symptoms. Maximum score is 21 with the higher the score, the greater the manifest anxiety.

^a Corpsmen in the same program but not participating in the study served as the matched sample. These Ss also received all the remaining tests, but in no instance did they vary significantly from those students actually selected for the study. This comparison of samples was undertaken in order to confirm the fact that Ss chosen were a random sample of corpsmen.

Intelligence. Intelligence was measured with the Army General Classification Test, (1945). Administered to entering military personnel, the test consists of two aptitudes, verbal (GCT) and mathematical (ARI). Intelligence represents the sum of the two subtests each given equal weight. In the population studied, the correlation between subtests was .74. The maximum intelligence value is 150.

Experiential time orientation. The Experiential Inventory (Cottle, 1968) explores predispositions to reflection (past orientation), momentary encounter (present orientation), and expectation (future orientation). Temporal designations are determined by listing important life experiences and assigning these experiences to a specific time zone.

The first half of the inventory's instructions are as follows:

Please list the 10 most important experiences of your life. These may be experiences you have had, you are having, and experiences you expect to have. You only need to write a few words for each experience and you may list your experiences in any order you wish.

At this point temporal orientations are implicit. In the second half of the instructions, however, time becomes the central operating variable:

Now that you have listed 10 experiences, please study the time zones below.

Time Zones

1. Distant Past
2. Near Past
3. Present
4. Near Future
5. Distant Future

Now take each experience and decide if it has occurred, is occurring, or will occur. Then choose the number of the time zone that best represents the time of the experience and write this number in front of the experience. Do this for all 10 experiences.

An experiential mean is calculated by summing the numbers of the listed time zones and dividing the number of experiences by 10. Essentially this value indicates where an individual locates himself in time: the higher the score, the more future oriented the S.

The mean is misleading, however, particularly in the center of the distribution of scores. For example, a mean of 3.0, theoretically a perfect present orientation, may result from locating all experiences in the present or from dividing experiences equally between the distant past and distant future. Thus, a numerical present orientation may mask an avoidance of present experiences.

To control for this, and for the purpose of establishing experiential orientation categories, respondents were divided into three groups. To a certain extent, the constitution of these groups is a function of the frequency distribution of experiences in the five time zones.

The first group consisted of those persons reporting

no near or distant future experiences. While they were called "past-present orienters" (pa-pr) they also may be thought of as being "future avoidant." In that all remaining persons listed at least one future experience, they were further divided into those listing three or more near and/or distant future experiences, so-called "future orienters," and those listing less than three but at least one future experience. This last group was labeled "samplers" as it turned out that they generally tended to distribute their 10 experiences throughout the five zones more widely than the other two groups.

It should be pointed out that while pa-pr orienters by definition list no future experiences, future orienters typically list some near and/or distant past experiences. This finding is anticipated somewhat by the rather low or pa-pr oriented experiential mean values (for males, $\bar{X} = 2.34$, $SD = .654$; for females, $\bar{X} = 2.38$, $SD = .624$).

Present duration. The Duration Inventory (Cottle, in press a) inquires into the temporal units selected for defining the duration of the past and future borders of the present, as well as the near and distant past and future.

Instructions for the Inventory are very simple:

Read the sentences below and fill in the blank spaces with one of the listed words. Indicate your choice by placing the number of the word in the blank space. Select only one word for each blank space.

1. Seconds; 2. Minutes; 3. Hours; 4. Days;
5. Weeks; 6. Months; 7. Years.

I. The PRESENT as I think of it extends from _____ ago to _____ from now.

II. The DISTANT PAST as I think of it includes things and events which occurred _____ ago, while the NEAR PAST includes things and events which occurred _____ ago.

III. The DISTANT FUTURE as I think of it, includes things and events which will occur _____ from now, while the NEAR FUTURE includes things and events which will occur _____ from now.

It should be noted that while an artificial distinction exists between clock (short) and calendar (long) units, a questionable scale incorporating the seven options nonetheless is established: the higher the number, the longer the duration. In this paper, only duration of the past and future borders of the present will be considered.

Temporal relatedness. The degree to which time zones are perceived as relating is examined in the Circles Test (Cottle, 1967). The Test's instructions are to

think of the past, present, and future as being in the shape of circles. Now arrange these circles in any way you want that best shows how you feel about the relationship of the past, present, and future. You may use different size circles.

Completed circle configurations normally conform to one of four designs:

1. Atomicity (Whitrow, 1961). Circles fail to touch and by demonstrating no spatial relatedness earn zero points on the dimension.
2. Continuity. By touching at peripheries only, circles depict a flow of units with each unit occupying its own finite "space." Two points are earned for each peripheral touch.
3. Integration (Murray, 1959; cf. Lessing, 1967; Wessman & Ricks, 1966). When circles overlap one another so that a particular space is shared by at least two zones, four points are earned.
4. Projection (Gurvitch, 1964). Circles may be drawn within (or projected into) one another or perfectly contiguous. That is, the space of the present and future may be identical. Either configuration receives a maximum of six points.

Total relatedness scores represent the sum of individual circle relationships: past-present, past-future, present-future. The summing operation poses a problem in interpretation, however, for as many as three different designs may be drawn in one configuration. In these instances assignment to a particular typology is made. The exact location of relatedness is determined by computing individual circle scores.

Other variables are generated in the Circles Test but they have been omitted in the present study (Cottle, 1967).

Temporal anxiety. The Temporal Anxiety (TA) scale consists of 12 statements dealing generally with the pace of life, fear of the future, and temporal orientations. The statements were randomly dispersed among a 39-item Time Attitude Inventory which instructed Ss to register agreement-disagreement on a 7-point scale.

To determine whether Ss judged time-anxiety items as belonging to a single cluster, a factor analysis of the 39 items was performed. A principal axis solution and Kaiser Varimax rotation operation was employed. Results of the analysis revealed all 12 anxiety items loading on the first factor. The factor consisted of the following items where the number in parentheses refers to the item's position in the Time Attitude Inventory: (32) I'm afraid I won't be able to lead a full life; (31) The pace of life is too fast for me; (19) I'm often so worried about what is going to happen, I forget about right now; (33) It's hard for me to work on a task when I know there's not enough time to finish; (34) I live in the past; (39) A minute seems too small an amount of time to be of any use; (23) I dislike change; (18) I live in the future; (20) More than knowing what's happening in the world, I like to know why things happen; (1) Sometimes I feel I'm racing time and losing; (11) Often I think how nice it would be to stop time; (15) I live in the present.

All the items load positively except for No. 15. The average loading value is 48.2 and the percentage of common variance is 30.4. Individual TA scale values were mechanically computed factor scores: High values indicate high temporal anxiety.

Future potency and evaluation. Potency and evaluation ratings were measured on Semantic Differential scales (Osgood & Suci, 1955; Osgood, Suci, &

Tannenbaum, 1957). The Semantic Differential instructs respondents to use adjective pairs for expressing, in this case, potency and evaluative attitudes toward the future. Each of the two factors consisted of five pairs of polar adjectives displayed on 7-point scales. The evaluation factor was comprised of the pairs: clean-dirty; good-bad; kind-cruel; fair-unfair; and pleasant-unpleasant. Potency adjectives included: hard-soft; large-small; strong-weak; brave-cowardly; and rugged-delicate. In each of the above pairs, the first word listed is the high-evaluation or potency word. In actual presentation, scales were intermixed with high words randomly assigned to the left- or right-hand side of the scale.

The presence of potency and evaluation factors was confirmed by factor analyzing results of the Semantic Differential. For this analysis, an orthogonal rotation was employed. In the present study, however, raw means of the two factors have been used.

Temporal fantasy. Interest in temporal recovery and preknowledge were measured on the Money Game (Cottle, in press b). In this task, fantasy is made explicit by transforming time allotments into purchasable commodities. Important to notice in the instructions is E's acknowledgment of the fantasy:

Pretend that you had a lot of money, more money than you could possibly use. Pretend also that someone had the power to sell you time, any time that you would want, and you would have this time given to you right now, knowing what you now know. Furthermore, having this time you could do whatever you wanted with it.

The Ss then are asked to consider the following money options which they may spend for the purchase of an hour, a day and a year from their personal past and future: (a) \$0.00, (b) \$10, (c) \$100, (d) \$1,000, (e) \$10,000. Interest is operationally defined as the amount selected. The Money Game also includes allotments from time before birth and following death, but these sections are not considered here (Cottle, in press b).

The various time allotments of the game were included as it was felt that while some persons might desire an hour's peek at their future, they might find a year too much or too threatening. Still others, eager to purchase a year might consider an hour too brief to be of worth. However, as the inter-correlations between allotments and payments were so high (average $r = .86$), the idea of examining individual allotments was dropped, and instead, the three payment amounts for each epoch were summed. Thus, fantasy scores ranged from no payments for any allotment to the full investment of \$30,000, \$10,000 for each allotment.

The issue of reality constraints has been purposely ignored in the Money Game. Nothing in the instructions states that obtaining premature glimpses of the future will in any way disturb "more normal" eventualities. Yet, the effect of money cannot be assessed as there is no way of knowing the degree to which playing with money influences the fantasy.

TABLE 1

CORRELATIONS OF TEMPORAL VARIABLES WITH VACH, MA SCALE, AND INTELLIGENCE

Temporal variables	Vach		MA scale		Intelligence	
	M	F	M	F	M	F
Experiential time orientation	.07	.19	-.04	.07	.06	.15
Temporal relatedness	.25**	.26*	-.26*	.11	.21**	.20*
Temporal anxiety	-.28**	-.28*	.55**	.48**	-.01	-.01
Present past duration	-.02	.18	-.02	-.07	-.10	-.09
Present future duration	-.02	.20*	.01	-.21*	-.13	-.18
Future potency	-.08	-.05	.07	.05	-.00	.15
Future evaluation	.06	.16	-.04	-.06	.10	.08
Past recovery	-.10	-.08	.26*	.12	-.11	-.09
Future preknowledge	-.07	.02	.01	.16	-.02	-.08

* $p < .05$.** $p < .01$.*Procedure*

Testing was done in group paper-and-pencil fashion during scheduled class periods under the supervision of a civilian examiner. Time instruments were administered in the second week, achievement and anxiety scales in the fourteenth week. GCT and ARI scores along with personal and family information were taken from Navy records.

To check stability of the time instruments, each was readministered in the twelfth week to all respondents. Reliability coefficients ranged from .73 on the Semantic Differential evaluation scale to .91 on the preknowledge section of the Money Game. Reliability values for the Vach scale and MA scale were not computed.⁴

RESULTS*Correlates of Achievement and Anxiety*

The correlations shown in Table 1 indicate that valuing achievement is positively correlated with temporal relatedness (Circles Test) and negatively correlated with temporal anxiety (TA scale). The prediction that achievement orientations would be associated with an extended future border of the present is found, however, only among women. Manifest anxiety on the other hand, as hypothesized, is inversely related to temporal relatedness among men only and positively related to TA scale scores of both sexes. The hypothesis that anxiety correlates with temporal fantasies is partially confirmed as men show

a correlation between temporal recovery and the MA scale ($r = .26, p < .05$).

In terms of simple correlations, intelligence relates significantly only to temporal relatedness (for males, $r = .21, p < .01$; for females, $r = .20, p < .05$). Thus, for the most part, performance on the time instruments occurs independently of intelligence. However, while remaining independent of anxiety, intelligence relates positively to valuing achievement. Achievement and anxiety are themselves negatively correlated. These findings are summarized in Table 2.

The Effect of Intelligence

The means and standard deviations of the temporal variables shown in Table 3 provide some indication of performance on the tests. In the cases of experiential orientations and temporal relatedness, the frequency distributions in the various categories are better performance indicators. These distributions are noted later.

Several findings in Table 3 deserve mention. First, women tend to expand both present borders more than men (past border $t = 2.54, p < .05$; future border $t = 1.34, ns$) who, in turn, show higher temporal recovery and preknowledge scores than women (recovery $t = 1.97, p < .10$; preknowledge $t = 5.60, p < .05$). Regarding this last finding, men reported that recovered time would be used for redoing such that present and future conditions might be altered. Thus time retrieved would be time remade. Women, however, re-

⁴ In addition to retesting, respondents were asked to describe their reactions to the temporal instruments and to tell what they had intended to communicate in the various tasks. In this manner, some measure of face validity was established.

TABLE 2
INTERCORRELATIONS OF VACH, MA SCALE, AND INTELLIGENCE

	IQ		Vach		MA scale	
	M	F	M	F	M	F
Intelligence	—	—	—	—	—	—
Vach	.24**	.23*	—	—	—	—
MA scale	-.04	-.07	-.14*	-.22*	—	—

* $p \leq .05$.** $p \leq .01$.

ported that time retrieved would be time relived.

The high TA scale scores demonstrate that anxiety about time is common among both sexes. In contrast, given the limitless wealth offered in the Money Game, fantasy scores are low.⁵ That payments for small allotments do not adequately account for the depressed summed means is demonstrated in the low payments for years alone (combined male-female past $\bar{X} = 2.81$; future $\bar{X} = 2.90$).

In order to test the remaining hypotheses and control for the effect of IQ, men and women were divided at the combined mean into high- and low-intelligence groups. Vach scores also were divided this way into high and low categories. Male MA scale scores

⁵ As a matter of interest, more than 70% of all respondents failed to pay anything for time of the historical past and future.

were trichotomized in order to locate possible curvilinear relationships. The smaller number of women in the sample, however, forced the use of a simple mean dichotomy.

Looking only at the statistically significant findings, an association is found between achievement value scores on the one hand, and experiential future orientations and low future potency ratings on the other. However, the former association is significant only among lower IQ men ($\chi^2 = 6.43$, $df = 1$, $p < .05$) and the latter only among higher IQ men ($\chi^2 = 4.58$, $df = 1$, $p < .05$). An additional discovery is that intelligence functions in raising the percentage of pa-pr orienters (35%) but in lowering the percentage of future orienters (36%) and samplers (29%).

High achievement among women tends to be associated with extended (i.e., above the mean) future borders of the present, but the

TABLE 3
MEANS AND STANDARD DEVIATIONS OF ALL VARIABLES

Variables	Males		Females	
	M	SD	M	SD
Independent				
Intelligence	113.4	11.72	109.6	10.22
Vach	5.62	1.11	5.41	1.03
MS Scale	8.73	4.14	9.39	4.09
Temporal				
Experiential time orientation	2.34	.654	2.38	.624
Past duration of present	2.92	2.30	3.51	2.46
Future duration of present	3.19	2.43	3.96	2.44
Temporal relatedness	2.61	3.56	2.12	3.21
TA scale	15.15	6.29	14.48	6.19
Future potency	5.12	1.03	5.11	1.15
Future evaluation	5.18	1.36	5.47	1.55
Past recovery*	6.84	1.62	6.38	1.58
Future preknowledge*	6.74	1.70	6.29	1.50

* Values summed for three allotments.

tendency reaches significant proportions only among the high-IQ group ($\chi^2 = 4.70$, $df = 1$, $p < .05$). Moreover, as the correlations between the variables suggested (Table 1), temporal relatedness is enhanced by valuing achievement independent of intelligence ($\chi^2 = 4.95$, $df = 2$, $p < .10$). Indeed, just high VACH women demonstrate integrated or projected designs which have been combined for the purpose of computing these chi-square values.

Whereas among lower intelligence men, both moderate and high anxiety levels are associated with temporal atomcity ($\chi^2 = 13.68$, $df = 4$, $p < .01$), among higher intelligence men, they are associated with high future evaluation scores, that is, scores above the mean ($\chi^2 = 5.68$, $df = 2$, $p < .10$), and high payments, that is, above the mean for time of the past ($\chi^2 = 8.32$, $df = 2$, $p < .02$) and future ($\chi^2 = 5.39$, $df = 2$, $p < .10$). Temporal anxiety, however, is related to manifest anxiety among low-intelligence men ($\chi^2 = 11.09$, $df = 2$, $p < .01$), high-intelligence men ($\chi^2 = 26.84$, $df = 2$, $p < .01$), and high-intelligence women ($\chi^2 = 6.24$, $df = 2$, $p < .01$).

A common characteristic of all the above findings is that where insignificant values obtain, the direction of the percentages normally is the same as in the significant cells. Hence, while intelligence in fact is sufficiently powerful to reduce associations to levels of statistical insignificance, in no instance does it reverse the effects of achievement and anxiety on time perceptions.

DISCUSSION

Results of the study may be summarized as follows: Achievement values in men and women encourage temporal relatedness on the Circles Test but function in lowering temporal anxiety (TA scale). Manifest anxiety, on the other hand, increases temporal anxiety and atomcity on the Circles Test. All of these findings occur independently of intelligence.

Among high-intelligence groups, achievement is associated with experiential future orientations in men and with extended future border durations of the present (Duration Inventory) in women. Manifest anxiety relates positively to evaluations of the future (Semantic Differential) and to increased interest

in recovering the past and foreknowing the future (Money Game). These last two findings, however, obtain for men only.

Among low-intelligence males, achievement functions in lowering potency and evaluation ratings of the future. Holding aside the effects of sex and intelligence, 8 of the 11 hypotheses are in some way confirmed. However, as anxiety failed to produce experiential past orientations, increase the present's past duration, or elevate perceptions of the future's potency, Hypotheses *c*, *d*, and *i* remain unconfirmed.

The universal finding that temporal anxiety increases with manifest anxiety but decreases through valuing achievement seems to reflect the dominant perceptions of the ways in which individuals incorporate the future. Whereas the achiever accepts the challenge of the unknown and turns the current of time into energy for activity and control, the anxious person finds the current constricting and overpowering, and ultimately seeks a return to the familiar and secure. Even with the variations attributed to intelligence levels, the image of the anxious man is one of overly swift time passage coupled with a desire to retrieve prior moments and foresee future ones. In addition, he appears unable to tack these moments onto present experiences. His world is one of clutching at a rushing time in order, perhaps, to still it or, at least, to keep a timeless past near him as long as possible: "Like a man on shipboard looking out over the stern at the foaming wake, he has turned his back on the future and sees time as a wave rushing backward into the past [Barrett, 1967, p. 362]."

The dilemma of turning back or facing forward has been articulated by psychologists and philosophers alike. Heidegger's description of the authentic man, for example, as one "whose basic possibilities have been disclosed to him including the possibilities of death, and who stands resolutely up to those possibilities [cited in Barrett, 1967, p. 361]" surely resembles descriptions of the achiever's special belief in personal control of outcomes (McClelland, 1961; Strodtbeck, 1958). These notions of possibility and personal control represent, moreover, the antithesis of Freud's (1936) notion that anxiety expresses helplessness, or Melges and Fourgerousse's (1966)

perhaps severe proclamation that "if an individual has no plans or cannot deal with the present in terms of the past, then behavioral disorganization will result [p. 138]."

This last passage, however, may help to explain the temporal relatedness variable and, for that matter, the synthesizing force inherent in the spread of the temporal horizon (Fraissee, 1963; Heidegger, 1962). If past-present connections are made through the sense of personal efficacy, control of activity or, more simply, autonomy, then present-future connections become the inferred extension of this autonomy. Inference gives life and meaning to the future. Prior achievement, therefore, breeds possibility and reinforces credible planning. Thus, indeed, "it is as if the [achievement] need has served to relate present achievement experiences to future ones [McClelland, 1951, p. 486]."

In philosophical terms, it is precisely in the action of planning that future possibilities or the sense of possibility is first confronted and ultimately actualized. But planning transcends feelings of competence, lists of significant expectations, or clinical admonitions, for according to Barrett (1967), "not only do we plan specific projects for today or tomorrow, but our life as a whole is a project in the sense that we are perpetually thrown-ahead-of-ourselves-toward-the-future [p. 361]."

Here, finally, is a key to understanding the characteristic responses of achieving and anxious persons to time. At one unequivocal level, time is only forward moving. We experience this as the sense of being-ahead-of-ourselves (Barrett, 1967; McClelland, 1961), as a forward look (Epley & Ricks, 1963), or future directedness (Kummel, 1966). Whereas the achiever accentuates this future directedness even to the point of seeking change and unfamiliarity (Winterbottom, 1958), anxiety leads one away from the future and thereby urges a magical undoing of an intractable reality. The undoing, which Freud (1936) associated with reactions to danger situations, may be the present tense and, therefore, the functioning mechanism of which temporal recovery (redoing) and preknowledge (pre-doing) are fantasy analogs. All three mechanisms fail, of course, as they either avoid or deny the linear flow of time with its irretriev-

able past, its unknowable future, and, most importantly, its finite limits of existence (Cohen, 1962, Ch. 5; Heidegger, 1962).

However, the sense of being-ahead-of-oneself must be qualified by the fact that men and women need not experience it in exactly the same way. The data indicate that valuing achievement affects frequencies of expectations and assessments of the *future* among men only. In women the value leads not to reverberations in future perceptions, but in future aspects of the *present*. Whereas masculine anticipation implies an active or agentic (Bakan, 1966) insertion into the future, the implication for women is an incorporation of now, or a continuous adding of moments to now. Achievement and anxiety, then, initiate systematic transformations of attitudes about the future but primarily among men (Bakan, 1966; McClelland, 1961), for, as a function of achievement, women lean more to a perception not unlike Gurvitch's (1964) "*present* [italics added] in advance of itself" (cf. Cottle, in press a).

Considering the sophistication of contemporary philosophical and psychological literature and particularly the stress on "the systematic description of what is given in experience [Barrett, 1967, p. 368]," commonly known as phenomenology, it is ironical that for the most part, definitions of past, present, and future remain unstudied. Future research must examine the kinds of primitive chronological differentiations made in the Duration Inventory, for these establish truly phenomenological temporal parameters. As they normally are conceived, temporal "orientations" become restricted in their applicability and even invalid if such dramatic variations are demonstrated in the determination of "simple" chronological "durations" (Cottle, in press a). Can one meaningfully discuss a future orientation, for example, if for some the future commences seconds from now, while for others an expanse literally of years appertains to what they themselves call present?

We must begin to speculate, furthermore, on the temporality inherent in social-psychological processes like achievement and anxiety, as both, in their way, manifest characteristic approaches to the future. Achievement strivings and anxiety are themselves experi-

ences of and in time, experiences which, presumably, the sexes do not equally share. And, while they remain negatively correlated in statistical operations, the complexity of that period of time known as the future reconciles many of their seeming disparities, both philosophical and psychological, as this passage reveals:

I am prepared for the unexpected, I lunge toward it with hunger, even though I've seen so much I seldom encounter a surprise; however, I am most near death when I begin believing I've seen it all. So I continuously doubt the future. My faith is behind me, not out in front. . . . I am pushing at crashing speeds into the unknown. . . . My power is in me, in all of us. Life is this power [Cartwright, 1967, p. 90]."

REFERENCES

- ARMY GENERAL CLASSIFICATION TEST. *Psychological Bulletin*, 1945, 42, 760-768.
- BAKAN, D. *The duality of human existence*. Chicago: Rand McNally, 1966.
- BARRETT, W. The flow of time. In R. M. Gale, *The philosophy of time*. Garden City, N. Y.: Doubleday, 1967.
- CARTWRIGHT, L. *To make a difference*. New York: Harper & Row, 1967. (Reviewed in *Time*, June 9, 1967.)
- COHEN, J. *Humanistic psychology*. New York: Collier, 1962.
- COTTLE, T. J. The circles test: An investigation of perceptions of temporal relatedness and dominance. *Journal of Projective Techniques and Personality Assessment*, 1967, 31, 58-71.
- COTTLE, T. J. The location of experience: A manifest time orientation. *Acta Psychologica*, 1968, 28, 129-149.
- COTTLE, T. J. The duration inventory: Subjective extension of temporal zones. *Acta Psychologica*, in press. (a)
- COTTLE, T. J. The money game: Notes on fantasies of temporal recovery and preknowledge. *Diogenes*, in press. (b)
- EPLEY, D., & RICKS, D. R. Foresight and hindsight in the TAT. *Journal of Projective Techniques and Personality Assessment*, 1963, 27, 51-59.
- FRANKE, P. *The psychology of time*. New York: Harper & Row, 1963.
- FREUD, S. *The problem of anxiety*. New York: Norton, 1936.
- GURVITCH, G. *The spectrum of social time*. Dordrecht, Holland: Reidel, 1964.
- HEIDEGGER, M. *Being and time*. (Trans. by J. Macquarrie & E. Robinson) New York: Harper & Row, 1962.
- HENRY, W. The business executive: The psychodynamics of a social role. *American Journal of Sociology*, 1949, 54, 286-291.
- KUMMEL, F. Time as succession and the problem of duration. In J. T. Fraser (Ed.), *The voices of time*. New York: George Braziller, 1966.
- LAVIN, D. *The prediction of academic performance*. New York: Russell Sage Foundation, 1965.
- LESSING, E. Developmental, personality and other correlates of length of future time perspective (FTP). Unpublished Research Report, Institute for Juvenile Research, 1967, 4, No. 2, 1-34.
- MACKINNON, D. W. A topological analysis of anxiety. *Character and Personality*, 1944, 12, 163-176.
- MAY, R. *The meaning of anxiety*. New York: Ronald Press, 1950.
- MCCLELLAND, D. C. *Personality*. New York: William Sloane, 1951.
- MCCLELLAND, D. C. *The achieving society*. Princeton, N. J.: Van Nostrand, 1961.
- MELGES, F. T., & FOURGEROUSSE, C. E. Time sense, emotions and acute mental illness. *Journal of Psychiatric Research*, 1966, 4, 127-139.
- MULLER, J. Self, time and activity. Unpublished manuscript, Department of Social Relations, Harvard University, 1967.
- MURRAY, H. A. Preparations for the scaffold of a comprehensive system. In S. Koch (Ed.), *Psychology: A study of science*. Vol. 3. New York: McGraw-Hill, 1959.
- OSGOOD, C., & SUCI, G. J. Factor analysis of meaning. *Journal of Experimental Psychology*, 1955, 50, 325-338.
- OSGOOD, C., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning*. Urbana: University of Illinois Press, 1957.
- PERSKY, H., GRINKER, R. R., MIRSKY, A., & GAMM, S. K. Life situations, emotions and excretions of hippuric acid in anxiety states. *Research Publication Association Nervous and Mental Diseases*, 1950, 29, 297-306.
- STRODTBECK, F. L. Family interaction values and achievement. In D. C. McClelland, A. L. Baldwin, U. Bronfenbrenner, & F. L. Strodtbeck (Eds.), *Talent and society*. Princeton, N. J.: Van Nostrand, 1958.
- TAYLOR, J. A personality scale of manifest anxiety. *Journal of Abnormal and Social Psychology*, 1953, 48, 285-290.
- TAYLOR, J. Theory and manifest anxiety. *Psychological Bulletin*, 1956, 53, 303-320.
- WALLACE, M., & RABIN, A. I. Temporal experience. *Psychological Bulletin*, 1960, 57, 213-236.
- WESSMAN, A. E., & RICKS, D. F. *Mood and personality*. New York: Holt, Rinehart & Winston, 1966.
- WHITE, R. Motivation reconsidered: The concept of competence. *Psychological Review*, 1959, 66, 297-333.
- WHITROW, G. J. *The natural philosophy of time*. New York: Harper & Row, 1961.
- WINTERBOTTOM, M. The relation of need for achievement to learning experiences in independence and mastery. In J. W. Atkinson (Ed.), *Motives in fantasy, action and society*. Princeton, N. J.: Van Nostrand, 1958.
- WOHLFORD, P. Determinants of extension of personal time. Unpublished doctoral dissertation, Duke University, 1964.

(Received August 30, 1968)

DIFFERENT PSYCHOLOGICAL EFFECTS OF LATERALIZED BRAIN DAMAGE¹

OSCAR A. PARSONS,² ARTHUR VEGA, JR.³

University of Oklahoma Medical Center

AND JULIAN BURN

Danville Veterans Administration Hospital, Illinois

Conflicting evidence exists concerning the differential psychological effects of lateralized brain damage on the Wechsler scales. Consideration of various reasons for discrepant results led to the selection of the WAIS Vocabulary and Block Design subtests as the best measures of verbal (language) and perceptual organization (visuoconstructive) factors. Comparison of mean scores of these subtests in unilateral lesion, bilateral lesion, and control groups of two separate experiments confirm hypothesized differences: right-hemisphere damage results in impaired visuoconstructive performance; left-hemisphere damage affects language abilities. The results could not be accounted for by variables such as age, education, and emotional disturbance, or localization, diagnosis, acuteness, and severity of brain disorder.

Behavior suggestive of a functional asymmetry of the two cerebral hemispheres in man has been the subject of numerous studies in the past 100 years. Accumulated clinical evidence has suggested that the left hemisphere is of primary importance in the control of language and symbolic processes, while the control of nonverbal, visuoconstructive abilities appears to be localized in the right hemisphere (Millikan & Darley, 1967).

Much of the experimental and clinical work leading to these hypotheses has involved the study of patients with unilateral cerebral damage. When quantification of verbal and nonverbal abilities and deficits has been attempted, one of the most frequently used instruments has been either the Wechsler-Bellevue or Wechsler Adult Intelligence Scale (WAIS). A number of studies have reported that left-hemisphere damage is accompanied by a decrement in Verbal IQ (VIQ), while

right-hemisphere dysfunction often results in a lowered Performance IQ (PIQ) (Guertin, Ladd, Frank, Rabin, & Hiester, 1966). Reitan (1962), in particular, has stated that such patterns of impairment are found regardless of the specific location of the lesion within the hemisphere.

Such findings, however, are by no means universal since equivocal or contradictory results have been reported, particularly with respect to PIQ. As Benton (1962) states,

The prediction that patients with lesions of the left hemisphere will show lower verbal scale scores than patients with lesions of the right hemisphere is hardly an exciting one, since we have been aware of this fact for decades. It is the prediction that patients with lesions of the right hemisphere will show greater deficit on non-verbal spatial tasks than will patients with lesions of the left hemisphere which is of particular interest. But it is precisely with respect to this prediction that the findings of the psychometric studies have been weakest [p. 258].

Smith (1966a, 1966b) probably represents the most outspoken critic of the notion of the differential organization of verbal and nonverbal abilities, at least as measured by the Wechsler scales. Studying rather large populations of brain-damaged Ss, he has failed to support the claim of differential effects of lateralized lesions, and suggests that sampling or population differences may be responsible for the findings of Reitan and his colleagues. Satz, Richard, and Daniels (1967) using

¹ This research was supported in part by National Institute of Neurological Diseases and Blindness Grant 05359. The authors are grateful for the cooperation of staff and patients of the University Hospital and the Oklahoma City Veterans Administration Hospital.

² Requests for reprints should be sent to Oscar A. Parsons, University of Oklahoma Medical Center, 800 Northeast Thirteenth Street, Oklahoma City, Oklahoma 73104.

³ Now at Danville Veterans Administration Hospital, Illinois.

multivariate covariance analysis to adjust for age, education, and intelligence, reported differential VIQ decrements in Ss with left-hemisphere lesions, but no differential impairment in PIQ of Ss with right-hemisphere lesions when these groups were compared with other populations. These authors also stress the importance of sampling differences.

When these negative findings with the WAIS are considered in relation to the many positive indications of the role of the right hemisphere in perceptual and visuoconstructive functions (Bogen & Gazzaniga, 1965; Teuber, 1962), it would appear that careful attention should be given not only to the characteristics of the population studied, but also to the nature of the dependent variable, that is, the Performance scale.

Benton (1962) and Heilbrun (1956) have called attention to the fact that the Wechsler Performance subtests are not truly "non-verbal," nor were they originally designed to be. To the extent that language and symbolic processes contribute to adequacy of performance on these subtests, patients with verbal-symbolic deficits might perform poorly whether or not they are displaying visuoconstructive deficits. There are many indications that the Performance scale is comprised of a more heterogeneous group of subtests than the Verbal scale. For example, the average correlation of subtests of the Performance scale with the total Performance score is lower than the comparable relationship on the Verbal scale (Wechsler, 1958). Further, Performance subtests do not have as high loadings on the nonverbal factor as do Verbal subtests on the verbal factor (Berger, Bernstein, Klein, Cohen, & Lucas, 1964). In short, compared to the Verbal scale, the Performance scale is not as good a criterion measure for testing the hypothesis of differential hemispheric control of psychological functions.

A more straightforward approach to the analysis of the differential effect of left- versus right-hemisphere pathology on verbal and what Guertin et al. (1966) term "perceptual organization" abilities would be to examine those WAIS subtests which best reflect these abilities. The results of a number of factor-analytic studies of the Wechsler scales have consistently shown that the Vocabulary and

Block Design subtests have the highest loadings (respectively) on the verbal and perceptual organization factors (Berger et al., 1964). Further, these subtests have the highest correlations with the Verbal and Performance IQs and are generally the most frequently chosen subtests when the full WAIS is not administered. For these reasons, the Vocabulary and Block Design subtests were used in the present studies. The general hypothesis to be investigated was that left-hemisphere damage leads to impaired language performance (Vocabulary score) and right-hemisphere damage results in impaired visuoconstructive performance (Block Design score).

An important methodological variable which has been emphasized by a number of investigators is that of the comparison groups. Hospitalized, non-brain-damaged patients frequently manifest a lower PIQ than VIQ (Guertin et al., 1966; Parsons & Kemp, 1960). Accordingly, inferences about the effect of damage to the right hemisphere on Performance subtests must take these findings into consideration.

Still another aspect of the sampling problem is that of equivalence of the severity and acuteness of the unilateral lesion groups. Arrigoni and De Renzi (1964) note that right-hemisphere cases are usually more severely brain-damaged than left-hemisphere cases because the symptoms of the latter lead to earlier identification. This suggests that severity of disorder is an important variable to be considered before inferences of differential impairment can be made. Similarly, acuteness or recency of brain disorder may lead to greater differential psychological impairment than long-standing or chronic conditions (Reitan, 1962).

With the above considerations in mind, two populations of brain-damaged patients from different experiments were separated into subgroups with lateralized damage in order to test the following specific hypotheses:

1. Within groups: patients with left-hemisphere damage will perform more poorly on the Vocabulary than on the Block Design subtests and the right-hemisphere-damaged group will manifest the opposite pattern.

2. Between groups: the patients with left-hemisphere damage will perform more poorly on Vocabulary than the right-hemisphere-damaged group, and the latter will perform at a lower level on the Block Design than the left-hemisphere-damaged group.

In each experiment, for comparative purposes, a group of bilateral and diffuse brain-damaged Ss and a group of non-brain-damaged, but hospitalized, patients were included.

METHOD

Data were collected in two experiments. The first experiment served as an initial test of the hypothesis; the second and more extensive study provided a cross-validation.

Subjects

Experiment I. Forty-four male veteran Ss with definite diagnoses of brain damage were rated by a neurologist⁴ experienced in using research rating scales, and placed into three groups: 20 Ss with predominately left-hemisphere damage (LHD), 16 Ss with predominately right-hemisphere damage (RHD), and 8 Ss with bilateral or diffuse damage (BHD). Forty-five male Ss without medical evidence of brain damage (NBD) served as a comparison group. These Ss came from medical, surgical, and psychosomatic wards. There were no significant differences among any of the groups on age (overall $M = 45.5$ years) and education (overall grade $M = 10.2$). LHD and RHD groups did not differ in the percentage of Ss falling into four main classifications of neurological disease. The Ss were distributed as follows: Trauma, 23%; CVA, 42%; Tumors, 25%; and Degenerative disorders, 10%. Further, there was no significant difference between the locations of the lesions within the hemisphere for the two groups: in the LHD group, the number of times the frontal, temporal, parietal, and occipital lobes were involved was 17, 9, 10, and 1, respectively. Comparable data for the RHD group were 10, 5, 10, and 0. All Ss were right-handed.

Experiment II. Ninety-eight Ss with unequivocal evidence of brain damage were rated by neurologists according to the same criteria as used in Experiment I. Twenty-five Ss were placed in the RHD group; 24 Ss in the LHD group, and 49 Ss in the BHD group. The NBD group consisted of 50 Ss drawn from medical, surgical, and neurological wards. The groups were comparable on age (overall $M = 45.9$) and education (overall $M = 9.53$), with the exception of the LHD group whose mean age (52.0 years) was significantly higher than the RHD, BHD, and NBD groups. However, since age-corrected scaled scores were used in the analysis, this difference was not deemed critical. Again, as in Experiment I, LHD and RHD did not differ in percentage of Ss falling

into the four neurological disease categories. The Ss were distributed as follows: Trauma, 14%; CVA, 53%; Tumors, 22%; and Degenerative diseases, 11%. For the LHD group, the number of occurrences of frontal, temporal, parietal, and occipital lobe involvement was 15, 14, 13, and 3, respectively; for RHD, the comparable figures were 13, 12, 14, and 7. Approximately 25% of the Ss in each of the groups were female. The bulk of the Ss were right-handed; one S in the LHD, one S in RHD, and 8 Ss in the BHD were left-handed.

As noted earlier, the S characterization of emotional disturbance and recency and severity of brain disorder may also lead to a differential deficit. If not taken into account, they may obscure the hypothesized relationships. The role of these variables is examined in the Results section.

Tests

It is important to note that Ss had to be sufficiently intact and communicative to undertake a rather extensive battery of tests over a period of several hours.

In both experiments, Ss were given the Vocabulary and Block Design subtests of the WAIS. In Experiment I, these were the only subtests from the WAIS used, but they were administered in the context of a larger battery of psychological procedures in a concept-identification study. All Ss were tested by one E (J. B.). In Experiment II, the full WAIS was given by different but experienced Es. Age-corrected scores were used in all analyses and one-tailed t tests were used to test the hypotheses.

RESULTS

Within-Groups Analysis

In Table 1 and Figure 1, the mean age-corrected Vocabulary and Block Design scaled scores are presented for the various groups. Within LHD and RHD groups, differences between the mean scores are in the predicted direction and significant: for LHD, Vocabulary is significantly lower than Block Design; for RHD, the opposite holds true. Within the BHD groups there were no significant differences between Vocabulary and Block Design scores. In the control group of Experiment I, Vocabulary was significantly lower than Block Design, an unexpected finding that would raise a question concerning interpretation of the LHD results were it not for Experiment II, in which the control group displayed no difference between means on the two subtests. The latter finding is in accord with previously reported results for control Ss from this general population (Parsons, Morris, & Denny, 1963). Therefore, the Vocabulary-Block Design difference for controls in Experi-

⁴ The authors are appreciative of the contributions of Fay Myers to their research.

TABLE 1

MEAN VOCABULARY AND BLOCK DESIGN AGE-CORRECTED SCALED SCORES AND *t* TESTS FOR WITHIN-GROUP DIFFERENCES

Group	\bar{X} Vocabulary	\bar{X} Block Design	<i>t</i>
Experiment I			
LHD	6.80	8.45	2.24*
RHD	8.94	6.06	2.75**
BHD	7.25	7.12	.18
Control	8.87	10.24	3.31**
Experiment II			
LHD	5.80	7.62	1.75*
RHD	8.28	5.80	3.03**
BHD	7.12	6.53	1.18
Control	9.72	9.86	0.11

p* < .05.*p* < .01.

ment I is considered to be due to sampling error.

Between-Groups Analysis

Again, the results were similar for both experiments (Table 2). The predicted relationships were found: LHD scored significantly lower than RHD on Vocabulary while RHD scored significantly lower than LHD on Block Design. Comparisons of LHD and RHD with the BHD group were nonsignificant with one exception: in Experiment II, BHD scored significantly lower than the RHD group on Vocabulary.

Comparisons of brain-damaged groups with

TABLE 2

t TESTS FOR BETWEEN-GROUPS DIFFERENCES

Group	Vocabulary	Block Design
Experiment I		
LHD-RHD	1.82* (LHD)	2.34* (RHD)
LHD-BHD	0.32	0.95
RHD-BHD	1.10	0.80
LHD-Control	2.49** (LHD)	2.16* (LHD)
RHD-Control	0.07	5.19** (RHD)
BHD-Control	1.25	2.61** (BHD)
Experiment II		
LHD-RHD	2.20* (LHD)	1.90* (RHD)
LHD-BHD	1.13	1.21
RHD-BHD	2.15* (BHD)	0.83
LHD-Control	4.16** (LHD)	2.92** (LHD)
RHD-Control	1.88* (RHD)	5.49** (RHD)
BHD-Control	3.97** (BHD)	5.00** (BHD)

Note.—Groups in parentheses have the lower mean score. (Means are reported in Table 1.)

p* < .05.*p* < .01.

controls gave rise to results consistent with expectancies. The RHD group in both experiments received significantly lower scores than controls on Block Design and on Vocabulary in Experiment II but not on Vocabulary of Experiment I (the latter probably due to the unusually low Vocabulary score of controls in Experiment I). The LHD groups were significantly lower than controls on both Vocabulary and Block Design. The BHD groups were lower on three of the four comparisons with the controls, the exception again being the Vocabulary scores in Experiment I.

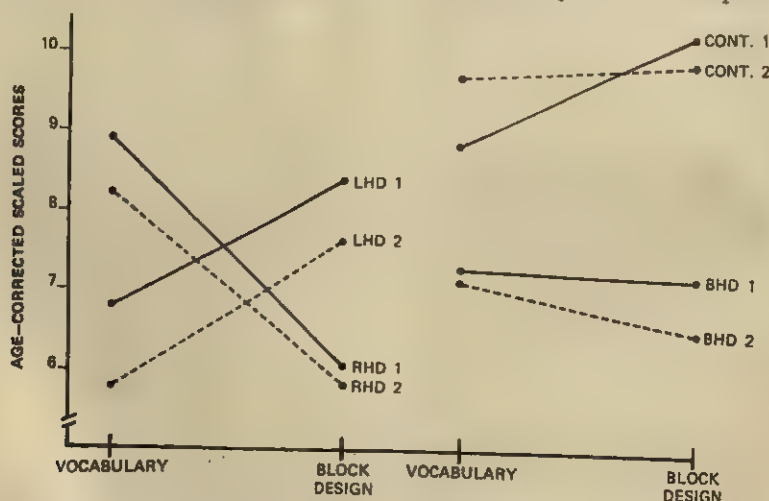


FIG. 1. Mean Vocabulary and Block Design age-corrected scaled scores for all groups in both experiments.

Correlational Analysis

Another approach to the hypothesized difference in hemispheric functions is possible through correlational analysis. If the hypothesis is valid, lower correlations between Vocabulary and Block Design would be expected in the unilateral lesion groups than in BHD and controls. In Table 3, Pearson product-moment correlations are presented for the two experiments separately and for the combined samples. It can be seen that correlations for the controls and RHD groups are remarkably consistent in the two experiments while correlations for the LHD and BHD are more variable. Using the combined sample as the best estimate of the correlation, RHD is the only group in whom the correlation is not significant. Tests for differences among these correlations indicate no significant difference among them, although the difference between RHD and BHD groups approached significance ($z = 1.80$, $p = .07$).

Additional Test of Hypothesis

In order to provide an additional test of the hypothesis, two other subtests of the WAIS, Similarities and Object Assembly, which have high loading on the verbal and perceptual organization (visuoconstructive) factors, respectively, were selected. The mean scores and correlations for these subtests for the Ss of Experiment II are presented in Table 4. One-tailed t tests for differences between means within groups and between groups were as follows: within groups between Similarities and Object Assembly, for LHD, $t = 2.52$, $p < .02$; for RHD, $t = 1.93$, $p < .05$. Between groups: LHD versus RHD on Similarities, $t = 2.48$, $p < .02$; on Object Assembly $t = 1.31$, $p = .10$. Thus the same significant relationships were obtained as reported above for Vocabulary and Block Design, with the exception of the nonsignificant difference on Object Assembly. Correlational analysis (Table 4) indicates that the RHD group has a significantly lower correlation between Similarities and Object Assembly, as tested by the r to z transformation (two-tailed test), than each of the other groups. The LHD group differed significantly only from the RHD group, and the BHD group differed significantly from RHD and controls.

TABLE 3

CORRELATIONS BETWEEN VOCABULARY AND BLOCK DESIGN AGE-CORRECTED SCALED SCORES

Group	Experiment I		Experiment II		Combined	
	<i>N</i>	<i>r</i>	<i>N</i>	<i>r</i>	<i>N</i>	<i>r</i>
LHD	20	.491*	24	.272	44	.369*
RHD	16	.203	25	.209	41	.235
BHD	8	.217	49	.581**	57	.549**
Controls	45	.465**	50	.461**	95	.442**

* $p < .05$.

** $p < .01$.

Other S Characteristics which May Have Influenced the Results

It is important to examine the possible effects of emotional disturbance, severity of psychological impairment, and recency of brain damage on the hypothesized differential impairments resulting from left- versus right-hemisphere damage. In several previous studies from our laboratories, hospitalized, non-brain-damaged Ss sampled in the same manner as in the present studies have had MMPI profiles similar to the brain-damaged Ss (Parsons et al., 1963; Parsons & Stewart, 1966). MMPIs were available only for the Ss of Experiment II. Since it is the RHD group in whom it is crucial to eliminate the possibility of the emotional disturbance per se giving rise to decrement on Block Design (Parsons et al., 1963), the MMPIs of this group were compared to those of the control group. The mean profile was similar for both groups: a heightened "neurotic triad" with secondary peaks on Pt and Sc . The mean T scores for the clinical scales (including Mf

TABLE 4

MEANS AND CORRELATIONS FOR SIMILARITIES AND OBJECT ASSEMBLY AGE-CORRECTED SCALE SCORES

Group	Similarities	Object Assembly	<i>r</i>
LHD	5.33	7.04	.527**
RHD	8.09	5.91	-.235
BHD	6.91	6.88	.657**
Controls	9.94	9.44	.310*

* $p < .05$.

** $p < .01$.

and S_i), a measure which correlates highly with degree of emotional disturbance (Parsons et al., 1963), were 58.8 for the RHD and 60.7 for the controls ($t = 1.00$, ns). Thus the RHD and control populations appeared to have similar psychological disturbances both in degree and kind. For the LHD the mean T score was 62, slightly (but not significantly) higher than the other two groups.

The two unilateral lesion groups were not significantly different as regards recency of brain disorder. Dividing the groups into those who tested within 6 months of onset and those who were tested after, there were 9 and 13 S_s , respectively, for RHD and LHD in the more recent damage groups. The average period of time between onset and testing for all S_s in both groups was 18.3 months. The pattern of differential impairment between LHD and RHD within the recent and chronic groups was the same as reported above for the overall analysis.

Finally, the important question of equivalence of severity of psychological impairment due to brain damage could be assessed in the LHD and RHD of this experiment. The Halstead Battery of neuropsychological tests had been administered to all S_s . Data have been reported which indicates that the Battery reflects presence and degree of severity of brain damage (Reitan, 1962). In our laboratory, the mean T score for all tests in the Battery has proved to be a potent discriminator between brain-damaged and non-brain-damaged populations (Vega & Parsons, 1967). The mean T score for the LHD was 37.9; for the RHD 40.0, a nonsignificant difference ($t = 1.14$).

It is concluded that the variables of emotional disturbance, severity, and recency of brain damage cannot account for the differential pattern of impairment observed in the LHD and RHD.

DISCUSSION

In two separate groups of patients with unilateral left- and right-hemisphere brain damage, differences in scores on the Vocabulary and Block Design subtests of the WAIS supported the hypothesis that left-hemisphere damage results in an impairment of language abilities and that right-hemisphere damage

results in a decrement in visuoconstructive ability. Further confirmation for the hypothesis was found in an additional analysis performed on the Similarities and Object Assembly scores of the S_s in Experiment II.

The consistent relatively clear-cut findings are attributable in part to the use of age-corrected scaled scores for each S . The well-known relationship between age and Performance subtest scoring is handled more effectively this way, in small samples of approximately the same age, than by the more traditional method of equating groups for mean age. Undoubtedly the use of the two WAIS subtests which best represent the hypothesized ability on the Performance scale, Block Design and Object Assembly, also contributed to the stability of the results. Support for this latter statement may be found in a comparison of results cited above (Experiment II) with findings based on IQs derived from using all subtests on each scale. Compared to RHD, LHD had a lower VIQ ($M = 89.1$ and 75.8 , $t = 2.79$, $p < .01$), but RHDs were not significantly different from LHDs on Performance IQ ($M = 78.7$ and 83.3 , $t = 1.13$, ns). Using the complete Performance scale attenuates differences in the RHD group and gives rise to nonsignificant results similar to those discussed by Benton (1962).

The correlational analyses provided additional evidence for differential hemispheric control of function. The correlations between tests representing a verbal factor and those representing the visuoconstructive factor in the RHD groups were consistently low and nonsignificant. In the LHD groups the correlations were significant and did not differ statistically from the controls or BHD. In accord with comments by Benton (1962) and Heilbrun (1956), it may well be that language factors contribute to a greater degree to performance on Performance scale subtests than the visuoconstructive abilities enter into the Verbal scale scoring.

There are at least three studies in the literature directly comparable to the present one in the problems to which they are addressed and in the use of Vocabulary or Similarities and Block Design or Object Assembly. The present results confirm those of Costa and Vaughan (1962) who found differences on

Vocabulary (Mill Hill) and WAIS Block Design as a function of lesion lateralization. The results also confirm the findings of Heilbrun (1956) in that Similarities was significantly lower in his LHD compared to RHD group. Heilbrun did not find overall significant differences between his lateralized lesion groups on a spatial battery of tests but did find differences on the Block Design when the severe dysphasics were removed from the analysis.

Finally, the findings provide partial confirmation of the results reported by Arrigoni and De Renzi (1964). These investigators, in a large-scale study of the effects of lateralized lesions, used the Verbal Scale of the Wechsler-Bellevue I and the Block Design and Object Assembly. They found the LHD group lower than RHD on the Verbal subtests, but the two groups were not significantly different on Block Design and Object Assembly. Arrigoni and De Renzi suggest that the positive results found by others could be due to the greater severity of damage in the RHD patients. Such a criticism does not hold for the present study in that the LHD and RHD groups of the major study (Experiment II) had the same level of impairment on the Halstead Battery. On the other hand, Arrigoni and De Renzi did not control for level of emotional disturbance in their Ss, a variable which has been shown to affect Performance scale subtests (Parsons & Kemp, 1960; Parsons et al., 1963).

It is clear that research in this area, especially that which attempts to demonstrate the role of the right hemisphere in visuoconstructive deficits, is fraught with difficulties. However, the present results, in which the variables of age, education, emotional disturbance and localization, diagnosis, severity, and recency of brain damage have been controlled or otherwise accounted for, give positive support to the growing body of evidence for the differential role of the hemispheres in psychological processes.

REFERENCES

- ARRIGONI, G., & DE RENZI, E. Constructional apraxia and hemispheric locus of lesion. *Cortex*, 1964, 1, 170-197.
- BENTON, A. Clinical symptomatology in right and left hemisphere lesions. In V. B. Mountcastle (Ed.),

- Interhemispheric relations and cerebral dominance*. Baltimore: Johns Hopkins Press, 1962.
- BERGER, L., BERNSTEIN, A., KLEIN, E., COHEN, J., & LUCAS, G. Effects of aging and pathology on the factorial structure of intelligence. *Journal of Consulting Psychology*, 1964, 28, 199-207.
- BOGEN, J. E., & GAZZANIGA, M. S. Cerebral commissurotomy in man: Minor hemisphere dominance for certain visuospatial functions. *Journal of Neurosurgery*, 1965, 23, 394-389.
- COSTA, L. D., & VAUGHAN, H. J. Performance of patients with lateralized cerebral lesions. Part I: Verbal and perceptual tests. *Journal of Nervous and Mental Disease*, 1962, 134, 162-168.
- GUERTIN, W. H., LADD, C. E., FRANK, G. H., RABIN, A. I., & HIESTER, D. S. Research with the Wechsler Intelligence Scales for Adults, 1960-1965. *Psychological Bulletin*, 1966, 66, 385-509.
- HEILBRUN, A. B. Psychological test performance as a function of lateral localization of cerebral lesion. *Journal of Comparative and Physiological Psychology*, 1956, 49, 10-14.
- MILLIKAN, C. H., & DARLEY, F. L. *Brain mechanisms underlying speech and language*. New York: Grune & Stratton, 1967.
- PARSONS, O. A., & KEMP, D. Intellectual functioning in temporal lobe epilepsy. *Journal of Consulting Psychology*, 1960, 24, 408-414.
- PARSONS, O. A., MORRIS, F., & DENNY, J. P. Agitation, anxiety, brain damage and perceptual-motor deficit. *Journal of Clinical Psychology*, 1963, 19, 267-271.
- PARSONS, O. A., & STEWART, K. D. Effects of supportive versus disinterested interviews on perceptual-motor performance in brain-damaged and neurotic patients. *Journal of Consulting Psychology*, 1966, 30, 260-266.
- REITAN, R. M. Psychological deficit. *Annual Review of Psychology*, 1962, 13, 415-444.
- SATZ, P., RICHARD, W., & DANIELS, A. The alteration of intellectual performance after lateralized brain injury. *Psychonomic Science*, 1967, 7, 369-370.
- SMITH, A. Certain hypothesized hemispheric differences in language and visual functions in human adults. *Cortex*, 1966, 2, 109-126. (a)
- SMITH, A. Verbal and nonverbal test performances of patients with "acute" lateralized brain lesions (tumors). *Journal of Nervous and Mental Disease*, 1966, 141, 517-523. (b)
- TEUBER, H. L. Effects of brain wounds implicating right or left hemisphere in man: Hemispheric interaction in vision, audition and somesthesia. In V. B. Mountcastle (Ed.), *Interhemispheric relations and cerebral dominance*. Baltimore: Johns Hopkins Press, 1962.
- VEGA, A., & PARSONS, O. A. Cross validation of the Halstead-Reitan tests for brain damage. *Journal of Consulting Psychology*, 1967, 31, 619-625.
- WECHSLER, D. *The measurement and appraisal of adult intelligence*. Baltimore: Williams & Wilkins, 1958.

(Received September 3, 1968)

RELATIONSHIP OF ACADEMIC UNDERACHIEVEMENT TO INCIDENTAL LEARNING

MICHELE S. MONDANI AND THOMAS A. TUTKO¹

San Jose State College

Forty-eight junior high school students were tested to determine whether or not there was any difference between achievers and underachievers in the learning of irrelevant material. All students were administered a test of social responsibility in a typical classroom situation. Subsequently, all were given a posttest on material printed on the responsibility test but irrelevant to it. Highly significant results on a *t* test between mean scores on the posttest of the achievers and underachievers demonstrated that underachievers learned significantly more of the incidental material. The problem for underachievement does not appear to be an incapacity to learn but a difficulty in focusing attention on the central material.

The concept of underachievement has been one that has gained a great deal of popularity in educational research. Studies have attempted to explore this multifaceted concept from many viewpoints. Numerous aspects of underachievement have been investigated including exploration of the underachiever who has superior ability (Miller, 1961), determining childhood factors (Winterbottom, 1953), describing and understanding the dynamics (Gowan, 1957), exploring frustration and hostility and its relationship to underachievement (Shaw & Black, 1960; Shaw & Grubb, 1958), and counseling and treatment of the underachiever (Miller, 1961).

Throughout these studies little attention has been given to the actual process of learning itself. Moreover, there has been no investigation of actual behaviors that an underachiever goes through during the exact time that learning should be taking place. Such a study should yield valuable insights not only in the understanding of the process by which underachievement takes place but also the means by which the problem may be at least partially alleviated. This study is an attempt to approach the problem from this very angle, that is, what behaviors are manifested by the underachiever during the learning period?

It is generally assumed that during the time in which learning is supposed to take place, the underachiever is learning nothing or at

least very little. One thing at least is readily apparent—he does not learn that which is required of him. But in fact it is hard to conceive of an individual learning nothing in a classroom situation. Although the underachiever may not be absorbing what the instructor intends for him, some learning ought to be taking place. He must therefore be “tuned in” to material which is irrelevant or incidental to the major learning material.

In an attempt to determine if this assumption is correct, the authors drew upon some literature from the area of experimental psychology called incidental learning which has been defined by Brown (1954) as “learning which occurs in the absence of (a) experimentally administered instructions to learn and (b) introspectively reported intention to learn [p. 163].” In essence, the major work done in this area indicates that individuals have a tendency to observe and learn materials which are incidental to a main body of knowledge. Usually these studies have employed research in which a main body of stimuli is presented to *S* and he must memorize it. In addition, other stimuli appear simultaneously but are considered irrelevant to the main learning task. If *S* remembers any amount of this “irrelevant” material, incidental learning is said to have taken place.

The major aim of this study was to determine if underachievers are learning the incidental information rather than the central information.

¹ Requests for reprints should be sent to Thomas A. Tutko, Department of Psychology, San Jose State College, San Jose, California 95114.

METHOD

Subjects

The 48 Ss (24 female and 24 male) employed in this study were all ninth-grade junior high school students.² All of the Ss were considered above average in ability and were placed in the most advanced two classes in the school. The school used a system of dividing the students into three main groups—X, Y, and Z. The X group was considered the high potential and performance group followed in rank order by the Y and Z groups. The students involved in this study were of the X and Y+ groups. It should be noted that while division of the students into groups is based on potential, performance and improvement are contributing factors as to whether or not the students remain in the original grouping. Hence, the Ss tested were those with the greatest potential and those whose performance proved adequate to remain in the more demanding class.

Although the students were performing sufficiently well to remain among the top-ranked groups, there were quite a few individuals labeled "underachievers." Thus, the students were divided into two groups—those who were performing below their predicted level and were subsequently considered underachievers and those who were performing at least at the predicted level. The evaluation and division of Ss into the two groups was made by teacher judgment. The teacher had known the students for almost a full academic year and divided them into either achievers or underachievers with consideration given to SCAT scores,³ past performance, GPA, and the general degree of effort demonstrated by the student. The use of teacher evaluation is a valid one, and, indeed, oftentimes the most reliable means of detecting such cases as latent underachievement (Miller, 1961). The Ss, then, were divided into two groups with 17 considered achieving below their ability (labeled underachievers) and 31 considered achieving at their predicted ability (labeled average achievers).

Procedure

The Ss were tested in a single group and every effort was made to maintain an atmosphere close to the average classroom situation. The Ss were told

² The authors gratefully acknowledge the cooperation and assistance of Peter Burnett Junior High School of the San Jose Unified School District.

³ Cooperative School and College Ability Test (SCAT). This test measures the capacity of students in grades 4-14 to undertake additional schooling. Four operational skills are measured: (a) getting the meaning of isolated sentences; (b) performing numerical computations rapidly; (c) associating meanings of isolated words; (d) solving arithmetic problems. The SCAT test is usually given in connection with the Sequential Tests of Educational Progress (STEP) which indicates the student's achievement while the SCAT indicates his ability.

they were to be part of a testing program and that all other information concerning the test and the purpose of it would be given at the conclusion of the session. They were given precise instructions on the use of IBM cards so that a normal emphasis was given to testing procedure as would be experienced with a task given by their usual instructor.

The Ss were then administered the Personality Scale for Social Responsibility in which there were 56 questions to be marked either true or false according to their personal attitudes (Gough, McClosky, & Meehl, 1952). Emphasis was placed in the instructions on the fact that there were no "right" or "wrong" answers and that honesty in answering was most important. In this way, each S set his own level of motivation since he was aware that in no way would he be graded.

This social responsibility test was only important insofar as it served as an orienting task; that is, it served to expose the Ss to the incidental stimuli. The incidental stimuli were as follows: Page 1—the date in the right-hand corner was May 1975; in the left-hand corner was typed D.C. 7; in the instructions the Ss were told to indicate a true response with a +; a line composed of question marks (????) separated the instructions from the body of the test; a star or asterisk (*) was placed beside Question 8. Page 2—the entire page was printed in green while the other pages were in purple ditto; there was a flower doodled in pencil as if it had been placed there by someone who had taken the test previously; Question 25 had another asterisk beside it. Page 3—a second doodle of a stick figure was placed at the top of the page; a comma was placed at the end of the last sentence on the page; the initials M.S.M. were placed at the bottom of the page. An effort was made to incorporate the incidental material in such a way so as to appear inconspicuous. For that reason, the doodles were drawn in pencil, the pages had been somewhat bent, and in general, the tests looked as if they had been previously used. Special and constant instructions were given not to communicate in any way with other students so as to minimize the possibility of any one student calling attention to the incidental material. The Ss were instructed to answer only the social responsibility test and to either look over the test or to turn it over when they had completed answering the questions. Once again, an attempt was made to maintain the typical classroom situation by permitting all students the same time with the material. All tests were collected at the same time and the Ss were again instructed not to talk.

Immediately following the collection of the responsibility test, a postexperimental test for the incidental material was administered. This postexperimental test consisted of 15 questions concerning the incidental stimuli. The questions were of both the recognition and recall types so that students were not handicapped or aided by their particular ability in test taking. A typical recall-type question was "There were two drawings on your test sheet (done in pencil). What were they?" An example of

a recognition-type question is "There was a line drawn between the instructions and the first question. Was the line made of: stars *, crosses +, question marks ?, dots ., or dashes -?" Again, Ss worked independently and were told there was no penalty for not remembering the material. The number of correct items on the incidental material formed the basis of comparison between the two groups.

RESULTS AND DISCUSSION

To determine whether there was a significant difference in the amount of incidental material retained by the achievers and the underachievers, a *t* test was run between the mean number of correct responses on the postexperimental test. The mean number correct for the underachievers was 9.29, for the achievers 5.77. The difference between the means was highly significant ($t = 3.16$, $df = 46$, $p < .01$). It is readily apparent that the underachievers retained a greater amount of incidental material than did those labeled achievers.

It would appear that there is a tendency of the underachiever to concentrate on those materials irrelevant to the learning task while the achiever concentrates on the material central to the learning task. In a sense the underachiever does well but only on the information which has little to do with the course content.

These results would be in line with the work being done with underachievers by Hewett (1965). In his hierarchy of tasks positively reinforcing the student for attention on the

course content is the basic step in having the student work toward greater achievement. Although the present study did not have actual classroom material as a learning task, it may be assumed to some degree that the findings in this study may be generalized to specific classroom learning material.

REFERENCES

- BROWN, G. H. Factors influencing incidental learning. *Journal of Experimental Psychology*, 1954, 48, 163-169.
- GOUGH, H. G., McCLOSKEY, H., & MEEHL, P. E. A personality scale for social responsibility. *Journal of Abnormal and Social Psychology*, 1952, 47, 73-80.
- GOWAN, J. C. Dynamics of underachievement of gifted students. *Exceptional Children*, 1957, 24, 98-101.
- HEWETT, F. A hierarchy of educational tasks for children with learning disorders. *Exceptional Children*, 1965, 31, 207-214.
- MILLER, L. M. *Guidance for the underachiever with superior ability*. Washington, D. C.: Government Printing Office, 1961.
- SHAW, M. C., & BLAKE, M. D. The reaction to frustration of bright high school underachievers. *California Journal of Educational Research*, 1960, 11, 120-124.
- SHAW, M. C., & GRUBB, J. Hostility and able high school underachievers. *Journal of Counseling Psychology*, 1958, 5, 263-266.
- WINTERBOTTOM, M. R. The relation of childhood training in independence to achievement motivation. University microfilm publication No. 5113, University of Michigan, 1953. Cited by L. M. Miller, *Guidance for the underachiever with superior ability*. Washington, D. C.: Government Printing Office, 1961.

(Received September 9, 1968)

SELF-CARE UNIT: SOME LESSONS IN INSTITUTIONAL POWER

ALAN P. TOWBIN¹

West Haven Veterans Administration Hospital, Connecticut

The hypothesis that the chronic psychiatric patient purposefully pursues a goal (continued hospitalization) that has harmful effects on him as a potential participant in the extrahospital community was the initial basis of a rehabilitation program. Initial failure led to a revision, which resulted in the return to the community of all 15 members. A framework for explaining the failure of the first and success of the second phase is presented: the staff's use of their power in defining the hospitalization situation. Exercise of power to redefine the hospitalization situation is probably a prerequisite to high discharge rates for chronic patients; redefinition of the conditions of admission is suggested as a basis for precluding the accretion of a chronic population.

In the recent past, traditional concepts of the chronic mental hospital inmate have undergone rapid change. The man once seen as the mentally debilitated victim of an irreversible disease process came to be regarded as the unwitting victim of institutional dehumanization (Goffman, 1961). More recently, the assumption that he is a purposeful participant in a social system, with goals that are deviant in relation to the system (Braginsky & Braginsky, 1967; Rakusin & Fierman, 1963; Simon, 1965; Towbin, 1966) has proven useful in the treatment of the chronic psychotic. These explanations may be integrated in the hypothesis that by his purposeful participation in the institutional system, the inmate may be harmed in the sense that dependency is fostered, ties to potential sources of support in the community are weakened, skills become obsolete, and he acquires a stigma (Scheff, 1966). This hypothesis informed the rehabilitation program reported here. In its initial form the program failed. A revision was instituted which proved very successful. In retrospect, the success of the revision indicates that modification of traditional assumptions about patients is not, in itself, an adequate basis for such a program. Traditional assumptions about the role of the staff, the conditions they create for the admission and discharge of inmates, and

especially their use of their power need to be scrutinized and modified. The two programs are described, and a theoretical analysis in terms of the staff's role in operating the system is used to explain the failure and success of the programs.

INITIAL PHASE

The staff of the chronic psychiatric wards of the hospital² had long felt that some modified living situation was needed to help patients make the move back into the community.

It was decided to set up a unit, explicitly identified as an experiment, to be evaluated at the end of 1 year. Laffal³ had established a self-care unit at this setting, but extraneous factors forced its closing before it could be assessed. The plan for the Phase 1 program was based on his prospectus. The rationale for the unit as set forth at that time was that

¹ Also affiliated with Yale University. Requests for reprints should be sent to Alan P. Towbin, who is now at the Learning Center, Farmington, Connecticut 06032.

² West Haven Veterans Administration Hospital, West Haven, Connecticut. The author is grateful to L. B. Fierman, Chief of Psychiatry, and J. Levine, Chief of Psychology, for their support of this project. Many psychology interns contributed to the project as members of the staff group. They include: M. Hersen, D. Travland, R. Hartsough, Suzanne Woolsey, Michaelene Plotnik, Arlene Brandon Marcus, K. Toney, D. Seelig, H. Kahn, and Fredda Kelly. Nursing staff members Gladys Punch, Edwina Diaz, and Sara Graham were involved throughout the two phases of the project. Grateful acknowledgment is hereby made to all of them. The original impetus for this project came from R. Almond.

³ J. Laffal. Prospectus for a self-care unit. Unpublished manuscript, 1964.

for certain patients, not necessarily all psychiatric patients, an interim setting of maximal independence and minimal ties with the hospital would be of therapeutic value in assisting the patient to adjust ultimately to complete independence from the hospital. These are patients who function adaptively in the institution, and do not seem to require complete hospital care but seem to be afraid to leave. Such patients seemed to have developed an adjustment to institutional living which is a handicap to independent living. We attempted to provide them with a setting in which they would experience intense social support through a high degree of peer-group interaction and as much autonomy as can be provided in a hospital. The plan for the unit was accepted by the chiefs of the services involved and the hospital administration. A physically distinct section of a chronic ward, a corridor of single and double rooms of 16 beds, was set aside for the unit.

In a prospectus shown to all applicants, a time limit of 6 months to 1 year was set forth; during the second 6 months, the progress of each member was to be reviewed the prospectus said, and all would be expected to leave the unit by the end of the year. Those unwilling or unable to leave the hospital by that time, could seek a transfer to another ward or institution, or reapply to the unit. Group therapy was offered to the unit (and never accepted) and individual therapy was made available where staff time permitted.

The following criteria were used for acceptance of members: that they be engaged in some regular activity in or out of the hospital—educational, prevocational, or vocational. Applicants who were not on interminable drug therapies were preferred. All applicants being treated with drugs picked up their drugs weekly at the pharmacy and took the medication as prescribed, as they would if living in the community. Applicants who were legally or Veterans Administration incompetent were rejected. They were advised on means of eliminating this bar to their acceptance. It was recognized that many applicants were attracted to the unit by the privileges of single or double rooms, freedom from staff intrusion, and extended sign-out, and could not muster a convincing statement that they

wanted to or hoped to leave the hospital. However, patients who gave this impression were not rejected if they were acceptable according to the explicit standards.

Members were told to organize themselves into a self-governing body with elected officers. This body was held responsible for all administrative functioning of the unit. After the first 10 or 12 members were selected by the staff, the staff passed on the eligibility of an applicant but the unit voted on whether or not to accept him. The unit was responsible for all decisions concerning discharge from the unit, disciplinary procedures for members, room assignments, reporting head count daily, as well as decisions concerning passes and leaves of absence. All members had the privilege of signing out *ad lib* between eight o'clock in the morning and midnight during the week, and for the entire weekend. Thus the unit had its own "turf" which required no contact with or supervision by nursing staff, although members could and did seek such contact.

After about 2 months, things seemed to be going well. There had been 17 patients on this unit. Three patients had been discharged from the unit since it opened—two to take jobs in distant communities, one to continue psychotherapy as an outpatient. Eight out of the 14 patients currently on the unit were working, in or out of the hospital. A group cohesiveness developed: patients wished one of their number to be called if a problem arose between nursing staff and a self-care patient. Since they took that stand, there had been no management problems with self-care patients.

SECOND PHASE

By the end of the year, however, a gloomier picture emerged. The *laissez-faire* policy had produced little result in terms of the hypothesis. The ward staff (nurses, nursing assistants, psychology trainees, and ward chief) spent several months reviewing the results and developing a plan to remedy the faults identified in the old one. It was felt that the patients had used the situation to continue their institutionally based lives with no action directed toward leaving. On the 4th of May, 1967, the second phase of the self-care unit (SCU)

was initiated. The following memorandum was sent to the chiefs of the psychiatry and the psychology services. Copies were presented to the members of the SCU as well.

The self-care unit was officially opened on February 11, 1966. A total of 22 patients had been on the unit in the following year. The outcome of this year is as follows: six were transferred to other wards, for the most part, back to G6W. Five of these 22 are on the unit now (and have been for six months or more) and are still employed. Five others are employed or in a full-time rehabilitation program. The remaining patients were discharged MHB; of these six, one is deceased, one has been readmitted following a period of employment: the remaining four are out of the hospital and employed full time. The "success" rate for this year is less than 20%. The policy of "laissez-faire" of setting no specific goals or standards, requiring no planning by SCU members for leaving the hospital, has not produced much in results. There has been some progress, especially in the increased responsibility of the members for each other, and the development of a more effective group spirit.

The members of the SCU have recognized that many of the patients on the unit have made no plans and are, in effect, idle. Their efforts to set norms for their associates has not been effective—perhaps because the staff has not participated enough. The staff of the ward has, therefore, decided to change the policy of the SCU as follows:

1. It is the aim of the SCU program to return patients to community living as quickly as possible. It is clear that a condition for community living for SCU members is a full-time job. We want to implement this policy as follows:
2. *The SCU group* is assigned the task of helping its members develop a plan for leaving the hospital.
3. This plan must have as its aim, obtaining a full-time job outside the hospital.
4. By two weeks from Monday each member will be expected to have come up with a detailed plan for discharge that is acceptable to the staff. No plan will be presented to the staff that has not been presented to the SCU group, and approved by the group. Any plan must include activity logically related to a discharge plan, and must be detailed with respect to the time the patient expects it to take.
5. The SCU is expected to make a weekly report on each member's progress—the development of his plans, and the steps he is taking toward realizing the plan.
6. Each member will have not more than 6 months to realize his plan, and leave the hospital.
7. Decisions about passes, leaves, etc., will be made by the SCU group, which should consider how these passes, leaves, etc., will contribute

to the realization of the member's discharge plan.

8. It is up to the group to make the necessary arrangements for implementing this policy—including increasing the frequency of meetings, if necessary.
9. The SCU as a group will be held responsible for carrying out this policy. Privileges of the group may be withheld if the group fails to carry out its responsibility to help each other prepare for a return to the community.
10. The ward staff will not meet with the SCU group, but invites the SCU to send a representative to a staff meeting, at which the progress report of the SCU will be reviewed. The time of this meeting will be announced. The ward staff will provide in writing an evaluation of the work of the SCU on its task of helping its members arrange to leave the hospital.
11. The SCU may, as a group, make recommendations to the staff, with respect to transfer, discharge, or other changes in the status of the members of the unit.

Many of the organizational features of this plan were taken from Fairweather's (1964) description of the "step-program" carried out under his direction. In the weeks that followed, there was a high rate of interaction between the unit and the staff. The requirements of the task were spelled out in more detail, and the unit was confronted with the assumptions which seemed implicit in the way they did not carry out the task. For example, the unit would present a member's medical complaints as grounds for that member's inactivity. Staff pointed out that experts (physicians) were required for determination of physical disability; if the unit needed expert opinions, staff would help to obtain them. This approach proved particularly effective in reducing the frequency of claims of physical disability. Similarly, it seemed implicit in the unit's messages to staff that members who claimed they had permanent and total disability for which hospital treatment could do nothing, expected to be boarded at this institution indefinitely. The unit was informed that the staff would help anyone who claimed that position find another place to live.

After about 1 month it seemed to the staff that members had united in opposition to the requirements of the program, and were not helping each other make plans to leave. They

were working together to help each other stay. The staff decided to lift the unique privilege of this group—that of staying out until midnight on their own cognizance. This was a blow to the prestige of the unit. It took about 2 months longer for the group to reach a degree of effectiveness in carrying out the task sufficient to move the staff to restore this privilege. The staff's weekly evaluation of the unit's discharge plans and progress reports maintained pressure on the unit to take action.

The unit seemed to respond by taking up each nonconforming member in turn and applying pressure, as a group, to get him going on a plan. Each member showed transitory "disturbance" but eventually complied. Fifteen patients were members of the unit at the end of June 1967. By November 3, 1967, 6 months from the inception of the second phase, 13 had been discharged, maximum hospital benefit. A month later, the fourteenth man left (23 days over his 6-month limit) and 2 months later, the fifteenth was discharged (within the 6-month limit for him).

Three of those who departed were "drop-outs" who left the hospital with little or no planning. The 12 graduates of the program include 8 with full-time jobs and a place to live. A ninth member, having reached the appropriate age, elected to retire. In keeping with his approved plan, he went from the hospital to the railroad station, where he boarded a train for Miami. Member 10 decided to live with his wife, remain a psychological invalid, and do the kind of volunteer work in the community that he had done in the hospital for many years. The last two men to leave had a place to live and were in full-time vocational rehabilitation programs.

A brief characterization of the SCU group follows. Of the 15, 10 had been in the unsuccessful Phase 1 program as well. Ages ranged from 28 to 60, with a mean 45 years. They had been hospitalized for between 3 and 77 months, with an average stay of 39 months. The range of Phase 2 stay was 63 to 206 days; the average, 121 days. Diagnoses include seven schizophrenic or other psychotic reactions, four depressive reactions, and four personality or psychophysiological disorders; four of the members had been diagnosed as having chronic brain syndromes as well. The

group was occupationally diverse: five unskilled laborers, one salesman, seven semi-skilled or skilled workers, one musician, and a nurse. Similarly their years of education ranged from 4 to 14, with an average of 10 years. Marriages, prior or current, are more frequent than is found in unselected populations of chronic psychiatric patients; seven were married, five separated or divorced, and three never married. While all 15 were discharged, three left without having an approved plan for their financial and living arrangements. Of the 12 graduates, seven sought outpatient follow-up treatment, and of these, five were discharged on psychotropic medication.

At this writing (6 months' follow-up)⁴ only one of the members is known to have applied to this facility for readmission. He was admitted, but returned to full-time employment after 1 month. But most striking to the staff was the fact that although the members had the option of making other arrangements for institutional living, they set a date for leaving in a planned way, and left. To members of the hospital staff who knew them, these patients appeared entrenched and intractable in their attachment to institutional living. And all had practical or realistic reasons for this attachment, in the sense that Simon (1965) has brought out; maybe they are crazy but they are not dumb.

Mr. A. was such a patient. He bitterly resented his wife, a well-educated woman, who seemed to delight in having him jailed for fracas that many lower-class families seem to tolerate without calling the police. He is almost illiterate, and claimed impotence in the face of all the forces arrayed against him by people who wanted to take advantage of him or who regarded him as a discredited person. He claimed he could not arrange a divorce; if he left the hospital and worked, his wife's

⁴ Using the procedure described by Fairweather, Simon, Gebard, Weingarten, Holland, Sanders, Stone, and Reahl (1960), 9 of the 12 "graduates" have been contacted for a 6-month follow-up. One of these had changed from vocational training to full-time employment status. Two, including the one member rehospitalized for a month, are described as getting along poorly, the rest satisfactorily. In other respects, the condition of the members appears about the same as at discharge.

creditors would gobble up his earnings. He had presented for admission with "paranoid" complaints about his fellow workers. Nevertheless he got his old job back, and left the hospital. He has since been divorced.

Mrs. B. is typical of the patients who use physical symptoms to disqualify themselves for independent living. She had been hospitalized for 5 years prior to the inception of the second phase of the SCU. Her prior psychiatric hospitalizations total an additional 8 months. While a patient on the locked service, she was grossly delusional and intermittently assaultive. Her entire hospital course was characterized by severe and persistent hypochondriasis, whiney complaining, tearfulness and dark hints that she knew the staff was keeping a diagnosis of cancer from her. She claimed that she could never work as a nurse again since her psychiatric hospitalization disqualified her, and her physical symptoms made regular work impossible.

Her position is of particular interest since she seems, in retrospect, to have maintained it by playing one staff member off against another, over the years. As a trained nurse, she elicited reports from the medical and surgical specialists summoned by her psychiatric resident for consultation on her recurrent complaints of back pain and constipation. She would then use the information obtained to try to discredit the ward nurse to the resident, and play on his fears that he might be overlooking something or neglecting her. The outcome of her interaction with the resident was almost invariably another round of consultation requests and revision of her medical regimen. When she attempted to go this route with the unit, the staff worked out a plan which was designed to prevent their being "split" by the patient (Stanton & Schwartz, 1954). Consultants were asked not to discuss their findings with her. After the neurosurgical consultant examined her back, the psychiatric resident told her authoritatively that there were no positive findings and that there was no objective basis for regarding her as physically disabled. This had the remarkable effect of appearing to relieve the patient. She dropped her physical complaints, perked up, found an apartment and a job as a nurse within a few weeks, and cheerfully departed.

She refused further outpatient psychotherapy on the grounds that she was feeling on top of things, though she continues to take a variety of medications.

DISCUSSION

The patient group was heterogeneous with respect to factors known to be related to length of hospitalization, successful rehabilitation, and low readmission rate. Yet, in terms of voluntary departure to a supportive social setting under tenable economic conditions all of the graduates must be considered successes. Long-term follow-up may reveal differences in adjustment which can be correlated with these prognostic indicators, but at this point the correlation is zero because the outcome has such low variability. The same can be said for "treatment" effects. Seven of the 12 graduates were treated with individual or family therapy during hospitalization, though two of these dropped out of treatment within 2 months of discharge. This highly regarded treatment is apparently unrelated to outcome in this sample. It seems that there must have been some factor (or factors) operating in this situation so powerful in its effect as to swamp the variability in the sample and produce a highly uniform outcome. The retrospective character of this analysis places severe limitations on the conclusions which can be drawn from it about the nature of the causal factors. With the understanding that the conclusions about what factors contributed to the success of the second phase are tentative and require more controlled study, they shall be set forth here.

The failure of the first and the success of the second phase of the SCU can be explained by the same set of factors if psychiatric hospitalization can be considered as a system with input, maintenance, and output phases. This system bears a relation to the rest of society comparable to that which the penalty box bears to the game of ice hockey. A player is removed from the game if he violates certain rules; he goes to the penalty box for a fixed period, leaving his team short one man. The referee makes an authoritative decision, based on his position and expertise, that a player's actions disqualify him. The penalty is of limited duration, but while he is out of the

game the player cannot influence it or the duration of his exclusion. He returns to the game automatically at the end of the penalty; his return to participant status calls for no authoritative intervention. Now, the citizen is transformed into an inmate by the authoritative opinion of an expert (psychiatrist), who decides that the citizen has a condition that indicates a need for hospitalization. The decision is based on expert assessment of the citizen's mental state, not on observable actions alone. Once the citizen is declared in need of hospitalization, he is assigned the status of "inpatient" and disqualified from participating in the game, so to speak. That is, he is disqualified from managing his own affairs. He is unable to influence the duration of his stay directly; efforts to shorten or prolong his stay are taken as mere manifestations of his condition, not requests to be granted or denied. Thus the patient is not only disqualified, he is discredited. Unlike the penalized skater, the former citizen can leave the patient status only if there is a decision complementary to the one that placed him in it. The expert, now responsible for the treatment of the patient's condition, must opine that the condition is sufficiently ameliorated that hospitalization is no longer needed.

This analysis leaves the status of the SCU member clouded. As was written in the prospectus, these were patients who no longer needed *complete* hospital care, but who seemed afraid to leave. What was not written, but believed, was that if we insisted on discharging them, they would do as they had done in the past: behave in ways that would make us regret our decision. The members of the unit were told they no longer needed hospitalization. The staff would appear to have effected the transformation of inmate back to citizen. But what we took away with the right hand, we gave back with the left. We ordered no discharge, urged no one to leave, kept no check of members' plans, and capped it all by the paternal gesture of providing a staff advisor to the unit. The staff had created an "experimental" situation, and we waited to see what effect it would have on what was defined as the crucial factor—the members' motivation to leave. As Towbin (1966) has noted, the clinical investigator is attracted to

the idea that he is a mere observer. But the members knew better. They knew that the staff has the power to order discharge for those who no longer need hospitalization, and the responsibility to use that power. Since the staff ordered no discharges, and was not trying to get anyone to leave, they were tacitly endorsing the members' stay. Motivation to remain in the hospital had been added to the staff's list of indications of continued need for hospitalization. We tacitly affirmed the existence of something we had explicitly denied—the members' need to be inpatients.

When we reviewed the failure of Phase 1, it was with some annoyance at the members for having taken advantage of us, and "holed-up" on the unit. Here we reacted much like the parents in Johnson's (1949) study, tacitly encouraging behavior about which we expressed dismay when the behavior could no longer be overlooked. Similar phenomena are described by Easson and Steinhilber (1961) and by Sargent (1962).

Of course, the staff did not carry out its work in this duplicitous way self-consciously and reflectively. At the time we planned the second phase we did have some sense that we had "let the patients get away with something" and decided to force the members to take a stand: return to the community or institutional living elsewhere. I believe that staff anxiety lest members act "crazy" or antisocially if discharged unwillingly, was probably ameliorated by the provision of this alternative. The experimental attitude was abandoned; their motivation to stay was no longer relevant. Departure of members was our goal. If anyone held out for institutional living, he could get it somewhere else, and for no better reason than that he preferred it. Authoritative support for the members' presence in the hospital, in the form of an identified indication of *need* for hospitalization, was withdrawn. There were no "escape clauses" in the second phase comparable to the vague "staff review" and provision that members may reapply if they haven't left in a year, provisions with which the Phase 1 prospectus abounded.

In the context of the traditional role of the expert, it might be said that we were merely providing an exit from inpatient status, by

declaring that members were not continuing to live in the hospital because they needed it. That is true, as far as it goes, but doesn't fully characterize the action. Up to 6 months more of living in the hospital was permitted, but members could choose to transfer to another institution without having to elicit an expert opinion that they needed it! Their preference would not be an indication, but a request to be granted if possible. No unit member took this alternative. Nor was the 6 months of permitted stay put forth as an interval during which the staff would "treat" members for their condition, so that they could qualify to leave. It was, on the contrary, a period during which they were required to carry out the task of making the transition from institution to community dweller.

The staff no longer related to the unit members in accord with traditional mental hospital practice. A voting membership of the staff group was established. This staff group administered the remainder of the ward by dealing with patients only via the task groups into which they had been organized. This format was then applied to the SCU. The unit members met with the staff as a group, and when the staff reviewed the unit's work and prepared an evaluation, a unit member was present as a representative.⁵

A consultant who was not part of the ward staff was offered to the unit to help them with their organizational problems, and was accepted by them. He never discussed his work with ward staff. The rate of interaction between unit and staff increased markedly. Unit members made more active use of the vocational rehabilitation services hitherto available but largely ignored. The idea that members were "sick" was actively discouraged in unit-staff communication, by the staff's substitution of "member" for "patient" in the memoranda. The members opposed this for a time,

always referring to themselves as "patient so-and-so" rather than mister or misses.

The staff's premise was that the secondary effects of hospitalization take time to overcome—loss of social contacts, drop in employability, etc., and that, therefore, we were not ready to discharge members immediately. This was the rationale for the task given the group, and the time to carry it out. In keeping with this task orientation, we frequently pointed out indolence and opposition on the part of the group, but never challenged their competence to carry out the task of planning their own lives. The goals of the staff, their standards and criteria, and the sanctions available to them were out in the open. It was not necessary to appeal to the staff's professional judgment or clinical experience or even the ultimate beneficence of our motives, let alone our superior judgment or wisdom in justifying our use of sanctions. We did not attempt to resolve conflicts between staff and unit by discrediting the members, and the complementary self-discrediting mentally sick role behavior of the members was not elicited.

The staff had shifted their role radically in the second phase of the SCU. Since this change is a departure from the traditional transformation and maintenance operations used as a framework for this analysis, it is easier to say what Phase 2 was *not* than to say what it was. Continued hospitalization had come to mean continued living in the hospital in Phase 2. It no longer expressed an expert opinion about the condition of a man's mind, nor were members regarded or related to as if they were disqualified or discredited. If anything, the 6-month limit expressed an opinion about the inmates' relation to the community, and an uneducated guess at the time necessary to carry out the task of changing that relationship. The staff were not treaters, the members not patients. The staff's definition of its responsibility changed from assessment of treatment needs and provision of a therapeutic milieu, to the assignment of a task and supervision of task performance. This included providing incentives for task-relevant behavior and deterrents to defiance. The status of SCU members was that of people who were temporarily in a living arrangement located in a hospital, while they did

⁵ This reorganization of the staff seemed to contribute to the staff's effectiveness. It provided opportunity for discussion of the strains created by the abandonment of traditional modes of relating to inmates, as well as providing new roles as participants in decision making about ward policy and action. This group activity seemed to enhance the staff's cohesiveness and morale, which were threatened by the massive opposition of the SCU members.

some work under expert (?) supervision. Exit from this status was defined by a time limit set by staff and the effectiveness of the members' performance. The benefits of hospital living were no longer available indefinitely, and had become untied from the type of self-discrediting behavior with which these people had hitherto affirmed their "need" for hospitalization. Indeed, those benefits under staff control were contingent on competence and effectiveness. The only basis on which the members could remain in the hospital was outright defiance of the staff; for example, a kind of sit-down strike might have succeeded in blocking the staff. The staff, it seems fair to say, had changed the definition of the situation: "hospitalization."

The highly selected nature of this group of chronic patients raises the question of the generality of these results. The writer has found that applying the same format to an unselected group of long-term psychiatric inpatients produced similar immediate results. That is, of 15 such patients 10 chose to leave the hospital from a program for which they had not volunteered, and did so in a planful, orderly way within a 6-month period. The other five chose transfer to other institutions. Unfortunately, no follow-up data has been obtained on these 10. Rothstein (1966) undertook a very similar program with larger numbers of chronic psychiatric inpatients, and reports on a 4-year follow-up. The program he describes also involves a staff-induced change in the definition of the hospitalization situation. Administration was individual, rather than via a patient-group as in the SCU. The conformity pressure exerted by the SCU group on its members, springing in part from the staff's power to lift the privileges of the entire group, seems to have had its counterpart in the staff's power to effect a "quick discharge" of any patient who refused to conform to the program or who openly defied it. The threat of prompt discharge induced most of the patients to expose themselves to a program of progressive responsibility and reward, leading to a self-chosen form of extrahospital living. Emphasis was placed on the exit phase, but the rules around admission were also changed so that if patients sought readmission, they would be required to reenter the

same discharge-oriented program. He reports:

Four years subsequent to discharge, 87 per cent of the patients were out of the hospital; the follow-up population as a whole averaged 83 per cent of time out of the hospital during the four-year post-discharge period; 50 per cent of these patients had spent no time in the hospital during this four-year period [p. 367].

In the framework of this analysis it seems that in both the Rothstein (1966) report and the program described here, the staff-produced redefinition of the situation of being hospitalized was the overriding factor leading to the choice of departure from the hospital. This redefinition was manifested in changes in the relations between staff and inmates in the implications of being in the hospital for one's qualifications and credibility, in the means available to influence the staff, as well as in the staff's use of its expertise. The conditions for maintenance of—and exit from—the patient role had been transformed. From this perspective, the chronic psychiatric patient may be defined as a person who, in pursuit of his own interests, has adapted to a situation as defined by the staff of the institution. The staff may be characterized as a group which usually behaves as if it were innocent of its power to define and redefine the situation.

The staff responsible for the programs reported here was grossly innocent of the radical changes it was imposing on a traditional mental hospital setting, and even more innocent of the power it wielded in doing so. One simply does not feel like a representative of society or the institution; one feels like oneself. We ignored the now obvious fact that power suffuses and colors all our actions in the mental hospital. An oriental sage said, "The fish does not know it lives in water, and does not care to have this pointed out." The mental health worker seems not to know he works in a power field. Many years ago Stanton and Schwartz (1954) called attention to this phenomenon.

And in spite of the political problems implicit in setting up a highly specialized type of government, the hospital attacked these problems handicapped by a tacit prohibition upon discussion of government, politics, power, influence, and sanctions [p. 244].

This innocence of power and the unwitting establishment of contingencies for access to

benefits is no less relevant to the input side of the hospital system than to the maintenance and output phases. The expert who makes the admissions decision is serving as gatekeeper to a very dense accumulation of societal resources. As Simon (1965) has noted, for those who are reluctant to leave the public mental hospital, these resources are nowhere else in as rich supply. Seen in terms of their use of power, the admitting staff sets standards of behavior (taken from psychiatric theory) as the condition of access to the resources accumulated in psychiatric hospitals. That is, the condition of access is described by the staff's concept of mental illness. At least for the applicants who become chronic inmates of mental hospitals, the power of the admission staff operates to elicit self-discrediting and disqualifying behavior (Braginsky & Braginsky, 1967). From a somewhat different perspective, Wood, Rakusin, and Morse (1965) make reference to this problem in their description of the conflict of the admitting resident.

The residents were alert to the fact that patients learn quickly that they would be admitted if they threaten violence. The possibility that they might be outwitted by a patient constantly plagued the admitting residents [p. 58].

As long as we operate in an atmosphere of innocence of our power, we can be very effective in innocently eliciting and fostering the behavior which confirms our theories and assumptions, and our professional role as well. One alternative is a radical transformation of the admissions situation, a redefinition comparable to that described here for the maintenance and exit phases of the mental patient's role. For example, applicants for admission would immediately be offered membership in self-care units, or some comparable organization, with the criterion of "admission" being their preference for an extracommunity living arrangement. Obviously, such a change in admissions practice implies far-reaching changes in the relation between the psychiatric inpatient facility and the community it serves, and does not lie within the power of the mental health professionals to bring about unilaterally.

CONCLUSIONS

The results of the programs reported here and the analytic framework used to explain their failure and success seem to be in keeping with a growing body of evidence and critical thought that the mental hospital as a system for the use of social power and the mental illness concept as the rationale for decision making in that system have serious harmful effects on the population they are meant to benefit (Ellsworth, Foster, Childers, Arthur, & Kroeker, 1968; Goffman, 1961; Kantor & Gelineau, 1965; Mendel, 1966). Many mental health professionals hope that the changes represented by the community mental health center approach will preclude or minimize the accretion of chronic inmates. This appears to be a forlorn hope as Kraft, Binner, and Dickey (1967) have shown in a 5-year study of turnover in such a center. Experience indicates that changing the physical setting, providing more "attention," and even abandoning traditional assumptions about the chronic mental hospital inmate, is not enough to increase discharge rates significantly. The traditional definition of the hospitalization situation, as established and maintained by the power of the staff, must be modified if the stasis of chronic populations is to change. The mental hospital is conceptualized as an organization for the distribution of societal resources; to the population served, psychiatric therapies are probably the least important benefits. Traditional admission practice makes self-discrediting behavior the ticket for access to these resources. A basic redefinition of the hospitalization situation would involve abandonment of self-discreditation as an access condition. Even if mental health professionals endorsed this change, it would undoubtedly be appreciated by the public as a fundamental change in the institution-community relationship, and is not likely to be accomplished without controversy.

REFERENCES

- BRAGINSKY, B. M., & BRAGINSKY, D. B. Schizophrenic patients in the psychiatric interview. *Journal of Consulting Psychology*, 1967, 31, 543-547.
- EASSEN, W. M., & STEINHILBER, R. M. Murderous aggression by children and adolescents. *Archives of General Psychiatry*, 1961, 4, 1-9.

- ELLSWORTH, R. B., FOSTER, L., CHILDERS, B., ARTHUR, G., & KROEKER, D. Hospital and community adjustment as perceived by psychiatric patients, their families, and staff. *Journal of Consulting and Clinical Psychology*, 1968, 32(5, Part 2).
- FAIRWEATHER, G. W., SIMON, R., GEBAND, M. E., WEINGARTEN, F., HOLLAND, J. L., SANDERS, R., STONE, G. B., & REAHL, J. E. Relative effectiveness of psychotherapeutic programs: A multicriteria comparison of four programs for three different patient groups. *Psychological Monographs*, 1960, 74(5, Whole No. 492).
- FAIRWEATHER, G. W. *Social psychology in treating mental illness*. New York: Wiley, 1964.
- GOFFMAN, E. *Asylums: Essays on the social situations of mental patients and other inmates*. Garden City, N. Y.: Doubleday, 1961.
- JOHNSON, A. Sanctions for super-ego lacunae of adolescents. In S. Glueck & E. Glueck (Eds.), *Searchlights on delinquency*. New York: International Universities Press, 1949.
- KANTOR, D., & GELINEAU, V. Social processes in support of chronic deviance. *International Journal of Social Psychiatry*, 1965, 11, 280-289.
- KRAFT, A. M., BINNER, P. R., & DICKEY, B. A. The community mental health program and the longer-stay patient. *Archives of General Psychiatry*, 1967, 16, 64-70.
- MENDEL, W. M. Effect of length of hospitalization on rate and quality of remission from acute psychiatric episodes. *Journal of Nervous and Mental Disease*, 1966, 149, 226-233.
- RAKUSIN, J. M., & FIERMAN, L. B. Five assumptions for treating chronic psychotics. *Journal of Hospital and Community Psychiatry*, 1963, 14, 140-148.
- ROTHSTEIN, C. Four-year follow-up of a non-traditional treatment program for chronic psychiatric patients. *Journal of Nervous and Mental Disease*, 1966, 142, 355-368.
- SARGENT, D. Children who kill—a family conspiracy. *Social Work*, 1962, 1, 35-42.
- SCHEFF, T. J. *Being mentally ill: A sociological theory*. Chicago: Aldine, 1966.
- SIMON, W. B. On reluctance to leave the public mental hospital. *Psychiatry*, 1965, 28, 145-156.
- STANTON, A. H., & SCHWARTZ, M. S. *The mental hospital*. New York: Basic Books, 1954.
- TOWBIN, A. P. Understanding the mentally deranged. *Journal of Existentialism*, 1966, 7, 63-83.
- WOOD, E. C., RAKUSIN, J. M., & MORSE, E. Resident psychiatrist in the admitting office: A man in conflict. *Archives of General Psychiatry*, 1965, 13, 54-61.

(Received September 9, 1968)

SENSATION SEEKING, IMPULSIVITY, AND PSYCHOPATHIC PERSONALITY

R. BLACKBURN¹

Broadmoor Hospital, Crowthorne, Berkshire, England

The hypothesis that sensation seeking is associated with impulsivity and psychopathic personality traits was investigated in a sample of psychiatric offenders. The Sensation Seeking Scale (SSS) was found to correlate positively and significantly with MMPI measures of impulsivity, extrapunitive, and psychopathy, and with scales of overt and covert hostility. However, diagnosed psychopaths did not score significantly higher on SSS than nonpsychopathic offenders. Evidence is discussed which suggests that while sensation seeking is related to psychopathic personality characteristics, it may also be common among certain psychotics. With this qualification, the results support the hypothesis.

Studies of exploratory behavior and the consequences of sensory deprivation in both animal and human Ss have indicated that living organisms are motivated by the absence or reduction of certain kinds of stimulus variability (Berlyne, 1960; Fiske & Maddi, 1961; Jones, 1966). An objective measure of the need for stimulus variability, the Sensation Seeking Scale (SSS) was developed by Zuckerman, Kolin, Price, and Doob (1964), and has been shown to predict premature quitting from and restless bodily movements in sensory deprivation and confinement experiments (Zuckerman, Persky, Hopkins, Murtaugh, Basu, & Schilling, 1966), and volunteering for unusual psychological studies (Zuckerman, Schultz, & Hopkins, 1967). Also, Farley and Farley (1967) found a significant positive correlation between SSS and the Extraversion scale of the Eysenck Personality Inventory, as predicted from the hypothesized relationship between extraversion and stimulus seeking (Eysenck, 1964). These studies go some way toward establishing the construct validity of the scale.

Evidence of a positive relationship between SSS and field independence as measured by the Embedded Figures Test led the authors to reject the possibility that the scale is a measure of impulsive or psychopathic characteristics (Zuckerman et al., 1964). However, subsequent findings cast doubt on this con-

clusion. Zuckerman et al. (1967) report a negative correlation between SSS and Gough's Socialization scale, and a small but significant positive correlation with the MMPI *Ma* (Hypomania) scale, which has been suggested to be a measure of impulsivity (Kassebaum, Couch, & Slater, 1959). Moreover, the relationship between extraversion and SSS would suggest that the latter is related to impulsivity since this is a major primary trait contributing to the introversion-extraversion dimension (Eysenck & Eysenck, 1963).

A need for excitement has long been considered to be a major characteristic of psychopathic personalities (e.g., McCord & McCord, 1964), and Quay (1965) has employed the concept of stimulus seeking to account for impulsive thrill seeking in such individuals. He proposes that much psychopathic behavior may represent a pathological need for sensory input, consequent on low physiological reactivity or rapid adaptation to stimulation, a proposition which has much in common with Eysenck's theory of the relationship between extraversion and antisocial behavior (Eysenck, 1964). Quay's hypothesis implies a positive relationship between stimulus seeking and impulsivity, and also that psychopathic individuals should show higher levels of stimulus seeking than nonpsychopaths. The present report concerns a test of these predictions. It was hypothesized that contrary to the earlier assertions of Zuckerman et al. (1964), SSS would correlate positively with questionnaire measures of impul-

¹ Requests for reprints should be sent to R. Blackburn, Department of Psychology, Broadmoor Hospital, Crowthorne, Berkshire, England.

sivity and psychopathic attributes, and that psychopaths would score more highly than nonpsychopaths on SSS.

METHOD

The Ss were male patients admitted to a maximum security hospital for psychiatric offenders over a 12-month period, were literate, able, and willing to complete psychological tests, and whose "cannot say" score on the MMPI did not exceed 50. The resulting sample consisted of 83 patients, about 70% of the total intake. Their mean age was 32.59 ($SD = 10.58$). Mean WAIS Full Scale IQ of 60 Ss who completed this test on admission was 99.53 ($SD = 11.99$). In terms of the 1959 Mental Health Act for England and Wales, 54 of the patients were classified as mentally ill (mostly schizophrenic) and 29 as psychopathic disorder. In clinical terms, however, only 21 of the latter group were subsequently diagnosed as psychopathic disorder or psychopathic personality.

The following tests were completed by patients within a week of admission: (a) the male form of the Sensation Seeking Scale (Zuckerman et al., 1964); (b) the Bendig Covert and Overt Hostility Scales (CH and OH) (Bendig, 1962); and (c) the MMPI, from which, as well as the regular scales, the following additional measures were scored: Welsh's Anxiety (A) and Repression (R) (Welsh, 1956); Gough's Impulsivity (*Im*) and Social Participation (*Sp*) (Dahlstrom & Welsh, 1960); Extraversion (*Ex*) (Giedt & Downing, 1961); the Hostility Battery (Foulds, Caine, & Creasy, 1960), which yields scores of General Hostility (GH) and Direction of Hostility (DH), positive scores on the latter indicating intropunitiveness. The various hostility scales were included in the investigation to measure antisocial tendencies. With the exception of Welsh's A, the remaining additional scales have all been found to have substantial loadings on the MMPI second factor, which has been identified as introversion-extraversion (Corah, 1964; Kassebaum et al., 1959).

RESULTS

The mean score of the sample on SSS was 9.94 ($SD = 4.44$), which is considerably lower than that of American students (Zuckerman et al., 1964) or British apprentices (Farley & Farley, 1967). However, this can probably be accounted for by the higher mean age of the present sample, since SSS correlated significantly with this variable ($r = -.333$, $p < .01$). The correlation of SSS with IQ was positive, but insignificant ($r = +.186$).

The relationship of SSS to personality and diagnostic measures is shown in Table 1. Only residualized coefficients are reported, age being partialled-out.

TABLE 1

CORRELATIONS OF THE SENSATION SEEKING SCALE WITH
MMPI AND HOSTILITY SCALES,
AGE PARTIALED-OUT

MMPI	Partial r
<i>L</i> (Lie)	-.256*
<i>F</i> (Validity)	.302**
<i>K</i> (Defensiveness)	-.238*
<i>Hs</i> (Hypochondriasis)	.051
<i>D</i> (Depression)	-.182
<i>Hy</i> (Hysteria)	-.041
<i>Pd</i> (Psychopathic Deviate)	.249*
<i>Mf</i> (Masculinity-Femininity)	.103
<i>Pa</i> (Paranoia)	.265*
<i>Pt</i> (Psychasthenia)	.135
<i>Sc</i> (Schizophrenia)	.222*
<i>Ma</i> (Hypomania)	.467***
<i>Si</i> (Social Introversion)	-.013
<i>A</i> (Anxiety)	.089
<i>R</i> (Repression)	-.359***
<i>Ex</i> (Extraversion)	.213
<i>Im</i> (Impulsivity)	.393***
<i>Sp</i> (Social Participation)	-.134
GH (General Hostility)	.258*
DH (Direction of Hostility)	-.389***
CH (Covert Hostility)	.251*
OH (Overt Hostility)	.283**

Note.— $N = 83$.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

The highest correlations are those with scales relating to impulsivity, namely *Ma*, *Im*, and *R* (Blackburn, 1968a; Kassebaum et al., 1959). Correlations with other scales relating to introversion-extraversion (*Ex*, *Sp*, *Si*) are in all cases not significant, suggesting that the correlation of SSS with extraversion (Farley & Farley, 1967) is carried largely by the variance attributable to the impulsivity rather than the sociability component of the introversion-extraversion dimension. Significant correlations are found with all the hostility measures, but more particularly with extrapunitiveness (DH) and with the *Pd* scale. All these correlations are consistent with the hypothesis that SSS is related to both impulsivity and to psychopathic characteristics.

Not predicted, but of some interest, are the correlations of SSS with *Pa* and *Sc*, and with the three validity scales (*L*, *F*, *K*). Together, these suggest a possible relationship between SSS and unconventional attitudes, and, perhaps, psychotic symptomatology.

The hypothesis predicts that patients diagnosed as psychopathic personalities will score more highly on SSS than nonpsychopaths. To test this, those patients diagnosed both legally and clinically as psychopaths were compared with 21 Ss drawn from the remainder of the sample, matched as closely as possible for age. (This matching was carried out "blind" by a colleague of the author, who did not know the identity or test scores of the Ss.) The difference between the psychopathic and nonpsychopathic offenders is in the predicted direction (Table 2) but does not reach an acceptable level of significance.

DISCUSSION

The correlations of SSS with the personality scales clearly indicate a fairly close association between sensation seeking and impulsivity, as predicted from Quay's hypothesis. The assertion of Zuckerman et al. (1964) that SSS does not measure impulsivity must therefore be rejected. Further, the significant correlations of SSS with *Pd*, extrapunitive (DH), and the scales of covert and overt hostility are consistent with the view that those who reveal psychopathic characteristics in self-report inventories will show relatively higher levels of stimulus seeking. On the other hand, the lack of a significant difference between diagnosed psychopaths and nonpsychopaths on SSS does not support the hypothesis. This result suggests two possible qualifications to Quay's hypothesis: either not all psychopaths show high levels of stimulation seeking and/or others with alternative diagnoses may also show such characteristics. There is evidence in the results relevant to each of these possibilities.

First, the criterion of psychopathy adopted in the present study was essentially the psychiatric diagnosis, and, as generally used by clinicians, the label refers to a somewhat heterogeneous group. Indeed, many of those described as psychopaths in the present sample were also labeled schizoid, hysterical, or inadequate, and it seems likely that by no means all of them would fit a more stringent definition of psychopathy such as that proposed by McCord and McCord (1964), that is, "an asocial, aggressive, highly impulsive person, who feels little or no guilt, and is

TABLE 2
COMPARISON OF PSYCHOPATHS AND
NONPSYCHOPATHS ON SSS

Group	N	Age		SSS	
		M	SD	M	SD
Psychopaths	21	28.81	8.18	11.62	4.77
Nonpsychopaths	21	29.00	8.37	10.29	4.24
<i>t</i>		.11		1.35*	

* $p < .1$, one-tailed test.

unable to form lasting bonds of affection with other human beings [p. 3].” It is, however, these very characteristics of asocial tendency, aggressiveness, impulsivity, and lack of guilt, as measured by personality scales, with which SSS is most significantly correlated in the present sample. An objective criterion of psychopathy based on a combination of these scales would clearly yield results consistent with the generalization that a high level of stimulation seeking is characteristic of the psychopath, but not perhaps characteristic of the more neurotic, guilt-ridden patient often labeled as an inadequate psychopath or character disorder (cf. Foulds, 1967).

However, the correlations of SSS with the *Pa*, *Sc*, and *F* scales of the MMPI suggest that sensation seeking may also be a feature of the behavior of some psychotic patients. There is evidence that among psychotic offenders, paranoid schizophrenics are both more extraverted and aggressive than non-paranoid schizophrenics (Blackburn, 1968b), and it is suggestive that of the three patients achieving the highest score on SSS in the present sample (a score of 20), two were diagnosed as psychopathic personalities, the third as a paranoid schizophrenic. In other words, it may well be that many of the attributes said to characterize the psychopath may also be common among paranoid schizophrenics. This being so, investigations of the possible causal antecedents of stimulus seeking proposed by Quay (1965) would best be approached by selecting Ss on the basis of personality characteristics of impulsiveness or aggressiveness, rather than in terms of diagnostic type.

REFERENCES

- BENDIG, A. W. Factor analytic scales of covert and overt hostility. *Journal of Consulting Psychology*, 1962, 26, 200.
- BERLYNE, D. E. *Conflict, arousal and curiosity*. New York: McGraw-Hill, 1960.
- BLACKBURN, R. The scores of Eysenck's criterion groups on some MMPI scales related to emotionality and extraversion. *British Journal of Social and Clinical Psychology*, 1968, 7, 3-12. (a)
- BLACKBURN, R. Emotionality, extraversion and aggression in paranoid and non-paranoid schizophrenic offenders. *British Journal of Psychiatry*, 1968, 114, 1301-1302. (b)
- CORAH, N. Neuroticism and extraversion in the MMPI: Empirical validation and exploration. *British Journal of Social and Clinical Psychology*, 1964, 3, 168-174.
- DAHLSTROM, W. G., & WELSH, G. S. *An MMPI handbook*. Minneapolis: University of Minnesota Press, 1960.
- EYSENCK, H. J., *Crime and personality*. London: Routledge & Kegan Paul, 1964.
- EYSENCK, S. B. G., & EYSENCK, H. J. On the dual nature of extraversion. *British Journal of Social and Clinical Psychology*, 1963, 2, 46-55.
- FARLEY, F., & FARLEY, S. V. Extraversion and stimulus-seeking motivation. *Journal of Consulting Psychology*, 1967, 31, 215-216.
- FISKE, D. W., & MADDI, S. R. *Functions of varied experience*. Homewood, Ill.: Dorsey, 1961.
- FOULDS, G. A. Some differences between neurotics and character disorders. *British Journal of Social and Clinical Psychology*, 1967, 6, 52-59.
- FOULDS, G. A., CAINE, T. M., & CREASY, M. A. Aspects of extra- and intro-punitive expression in mental illness. *Journal of Mental Science*, 1960, 106, 599-610.
- GIEDT, F. H., & DOWNING, L. An extraversion scale for the MMPI. *Journal of Clinical Psychology*, 1961, 17, 156-159.
- JONES, A. Information deprivation in humans. In B. Maher (Ed.), *Progress in experimental personality research*. Vol. 3. New York: Academic Press, 1966.
- KASSEBAUM, G. G., COUCH, A. S., & SLATER, P. E. The factorial dimensions of the MMPI. *Journal of Consulting Psychology*, 1959, 23, 226-236.
- McCORD, W., & McCORD, J. *The psychopath*. Princeton: Van Nostrand, 1964.
- QUAY, H. C. Psychopathic personality as pathological stimulation seeking. *American Journal of Psychiatry*, 1965, 122, 180-183.
- WELSH, G. S. Factor dimensions A and R. In G. S. Welsh & W. G. Dahlstrom (Eds.), *Basic readings on the MMPI in psychology and medicine*. Minneapolis: University of Minnesota Press, 1956.
- ZUCKERMAN, M., KOLIN, E. A., PRICE, L., & DOOB, I. Development of a sensation-seeking scale. *Journal of Consulting Psychology*, 1964, 28, 477-482.
- ZUCKERMAN, M., PERSKY, H., HOPKINS, T. R., MURTAUGH, T., BASU, G. K., & SCHILLING, M. Comparison of stress effects of perceptual and social isolation. *Archives of General Psychiatry*, 1966, 14, 356-365.
- ZUCKERMAN, M., SCHULTZ, D. P., & HOPKINS, T. R. Sensation seeking and volunteering for sensory deprivation and hypnosis experiments. *Journal of Consulting Psychology*, 1967, 31, 358-363.

(Received September 9, 1968)

PERSONALITY CHARACTERISTICS OF PARENTS OF NEUROTIC, AGGRESSIVE, AND NORMAL PREADOLESCENT BOYS¹

LORNA M. ANDERSON²

University of Minnesota

MMPI's of parents of neurotic, aggressive, and normal 9- to 11-year-old boys were compared, using two of Achenbach's factors for classification of child psychiatric patients. Twenty-three couples, parents of aggressive boys, were more deviant than 29 couples, parents of neurotic boys. Experimental parents were more deviant than 50 control couples, parents of normal boys. Fathers of aggressive boys were higher on *Pd* than either fathers of neurotic or control boys. Fathers of neurotic boys were higher on *K*. Mothers of aggressive boys had more low point *Mf*. More neurotic boys were oldest and more aggressive boys middle children.

Interest in parent personality patterns associated with behavior problems in children has increased as diagnosticians and therapists have tried to understand and treat other family members who may contribute to emotional disorders in children. Although there is considerable research involving parents and children, Butcher (1969) points out that most studies dealing with whole families have not been well controlled for important *S* variables. In addition, not many studies are comparable because of the wide variety of research methods used. Some conclusions can be drawn, however. Parents of disturbed children are more disturbed than nonclinic parents, but less deviant than psychiatrically disturbed adults. One of the most specific findings is that there appears to be a strong relationship between aggressive, openly punitive parental behavior and childhood aggression, particularly in father-son relationships. For example, Butcher (1966) found fathers of adolescent aggressive boys had significantly

higher elevations on the *Pd* scale of the MMPI.

A major problem in "whole family" research is that of classifying *Ss* in some meaningful way. One method of study is to categorize the children according to symptoms, traditional psychiatric diagnoses, or some other criteria, and then examine the parents to see if there are meaningful personality patterns associated with a given category in the child. The traditional psychiatric diagnoses do not satisfy the research requirements of objectivity and reproducibility, however. Therefore a different system of classification of child psychiatric patients is needed.

The externalizer-internalizer dimension for describing adult psychiatric patients proposed by Philips and Zigler (1961) was found reliable in classification of patients into "externalizers" who "act out" their problems, often in antisocial behavior, and "internalizers," who resort to fantasy. Achenbach (1966), working with childhood behavior problems, found an externalizer-internalizer dichotomy similar to that proposed by Philips and Zigler. Using clusters of symptoms, he was able to classify some boys, for example, as "aggressive" or "externalizers," who were characterized by antisocial, acting-out behavior; other boys could be classified as "obsessive-compulsive-phobic" or "internalizers."

This classification into externalizers and internalizers appears applicable to both the children and adults in families, and much existing research into the interaction of

¹This investigation was supported by a Public Health Service Fellowship, No. 1 F1 MH-25, 347-01, from the National Institute of Mental Health. It is based on a dissertation, submitted March 1968, to the Department of Psychology, University of Minnesota, in partial fulfillment of the requirements for a PhD degree. The author wishes to acknowledge the continued guidance, valuable suggestions, and editorial assistance received from James N. Butcher.

²Requests for reprints should be sent to Lorna M. Anderson, who is now at the Child Psychiatry Department, St. Paul-Ramsey Hospital, Saint Paul, Minnesota 55101.

parent and child personality could be viewed from the standpoint of these two categories. For example, the typical parent of a delinquent child might be expected to be an "externalizer" since he tends to act out by using physical punishment and behaving in a generally openly aggressive fashion (Bandura & Walters, 1959; Glueck & Glueck, 1950).

The purpose of this study was to examine relationships between parents' personality and specific behavioral problems in their children. The primary goal was to study personality characteristics of parents of two well-defined and homogeneous clinical populations, "aggressive" (A) and "neurotic" (N) boys, following Achenbach's factors (1966). Background data selected for study were age of parents, occupation and education of fathers, size of family, birth order of boy, and IQ of clinic boys.

METHOD

Subjects

Experimental Ss were parents of 9- to 11-year-old Caucasian boys who had been patients at four psychiatric facilities.³ Only those parents were included who were married, living together, and were natural parents of the boys. The MMPI had been administered to all of these parents as part of initial diagnostic procedures.

In addition, the boys whose parents were included in the study had to meet several other criteria. Their age was between 9 and 11 years inclusive at time of initial contact with the family. IQ had to be 90 or above, or in a few cases where no IQ was available, there was evidence of at least average grades in school. There was no evidence of brain damage or severe physical disabilities. Initial survey of the files resulted in a sample of 395 sets of parents whose sons met the above criteria. Of these, 29 pairs fit the experimental A classification, and 23 pairs were experimental N.

Control Ss were 50 pairs of parents of Caucasian boys in regular (not "special") fourth, fifth, and sixth grades.⁴ All these parents were married, living together, the natural parents of the boys, and presumably psychologically healthy.

Measures

Measures used in the study were the MMPI, a background data sheet for clinic families, a less de-

tailed background data sheet for control families, and a 91-symptom check list developed by Achenbach (1966). In addition, Achenbach's symptom lists for his "aggressive behavior" and "obsessive-compulsive-phobic" rotated factors were used to determine the classification of Ss—labeled here experimental A or aggressive and experimental N or neurotic, respectively. These categories were chosen rather than externalizers or internalizers because they appeared to represent the extreme end points of Achenbach's internalizer-externalizer first principal factor. It was hoped this restriction would minimize the chance of overlapping symptoms in the children, and thereby insure more distinct, homogeneous parent groups. (It should be noted all the aggressive boys in this study would also be classified as externalizers, and the neurotics as internalizers.)

MMPI scores and background data for experimental Ss were obtained from clinic files in which initial diagnostic work had been completed. The child's presenting complaints were checked off on the 91-symptom check list according to the procedure devised by Achenbach. Psychologists' and psychiatrists' reports were referred to only for intellectual evaluations and tests for organicity. If 60% of the symptoms coincided with experimental List A or List N, they were labeled accordingly. (The researcher was the sole judge, since previous studies showed an average interjudge reliability of 80%.) The control group of parents were all volunteers who completed the MMPI and background data sheet in supervised group testing. Each couple was also contacted individually in an effort to elicit any possible abnormalities in parent or child which might not meet conditions of selection.

RESULTS

Background data analysis (using a chi-square test of proportions) showed no significant differences except in birth order of the clinic child, with neurotic boys tending to be firstborn and aggressive boys middle children ($p < .01$).

The parent MMPI's were compared by a multivariate analysis of variance, an analysis of frequency of high and low points in individual *K*-corrected profiles, inspection of group mean profiles, and an analysis of configural differences between *Hy-Pd* and *Pd-Mf* scale scores. Individual MMPI's were excluded where scales *?* or *L* were greater or equal to a *T* score of 70, or where the *F* or *K* scales were greater or equal to a *T* score of 75. Mean *T* scores for the six groups are given in Table 1 and group mean profiles are shown in Figure 1.

A multivariate analysis of variance was performed using raw scores. (This method of

³ University of Minnesota Hospital Psychiatric Department, Institute of Child Development Clinic, Washburn Clinic, all in Minneapolis, Minnesota; and the Wilder Clinic, in St. Paul, Minnesota.

⁴ Pupils at the Hayes Elementary School in Fridley, Minnesota.

TABLE 1

MEAN MMPI T SCORES FOR EXPERIMENTAL A, EXPERIMENTAL N, AND CONTROL MOTHERS AND FATHERS

Scale	Experimental A		Experimental N		Control	
	Mothers	Fathers	Mothers	Fathers	Mothers	Fathers
<i>L</i>	50	46	50	50	50	46
<i>F</i>	53	53	50	50	50	50
<i>K</i>	55	53	59	61	59	55
<i>Hs</i>	56	54	52	52	52	50
<i>D</i>	55	56	57	51	51	53
<i>Hy</i>	61	56	59	60	56	56
<i>Pd</i>	57	64	56	57	55	50
<i>Mf</i>	43	54	47	57	47	57
<i>Pa</i>	59	53	56	53	56	53
<i>Pt</i>	56	54	55	52	56	54
<i>Sc</i>	57	55	54	51	54	51
<i>Ma</i>	53	55	45	53	50	55
<i>Si</i>	54	52	51	48	53	52

analysis did not permit corrections for *K* and sex differences included in the usual *T*-score profiles.) There were six cells with the 13

mean MMPI scale scores nested in 13×13 matrices within each cell, the final *F*s based on determinant ratios.

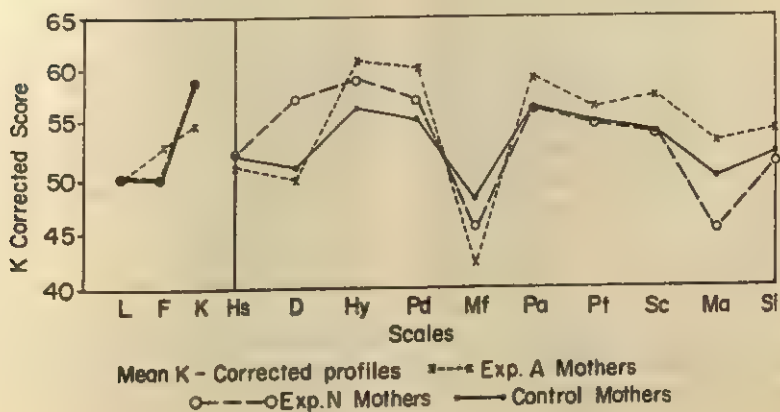
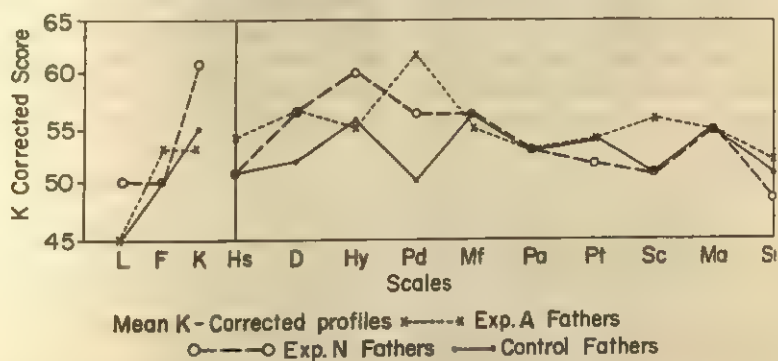


FIG. 1. Profiles of fathers and mothers.

TABLE 2

RESULTS OF MULTIVARIATE ANALYSIS OF VARIANCE OF
MEAN MMPI RAW SCORES, EXPERIMENTAL A,
EXPERIMENTAL N, AND CONTROL
MOTHERS AND FATHERS

Source	df	F value	p value
Grand mean	13/186	2299.12	<.000000001
Sex	13/186	39.09	<.000000001
Treatment ^a			
Controls vs.	13/186	3.07	<.001
Experimental			
Experimental A vs.			
Experimental N	13/186	2.75	<.01
Sex × Treatment	26/372	1.05	.40 (ns)

^a Overall treatment effect is average of two treatment analyses given.

In addition to the grand mean, sources of variation tested were sex, treatment, and the Sex × Treatment interaction. Treatment effects were broken into two parts: controls versus experimental N and A combined (i.e., controls versus experimental) and experimental N versus experimental A.

The results of the multivariate analysis are given in Table 2. Only the interaction effect, Sex × Treatment, was not significant.

Univariate ANOVAs

Univariate tests within treatment, sex, and Sex × Treatment main effects were also performed, as shown in Table 3. None of the interaction effects was significant.

Seven of the univariate tests for sex reached significance. Most of the variation resulted from using raw scores in the analysis, since no allowance had been made for the known differences in male and female *T* scores. On scales *Pa* and *Ma*, however, a significantly different response between males and females was found, where such differences are not normally expected. In the absence of significant interaction effects, it appeared this difference extended to all three groups, experimental and control both. Females were uniformly higher than males on the *Pa* scale and males were higher than females on the *Ma* scale.

Significant differences between the experimental groups combined versus controls were found on scales *Hs*, *D*, *Hy*, and *Pd*, with the

experimental group more deviant on each of these scales.

Significant treatment effects for experimental A versus experimental N parents were found on scales *K*, *Hs*, *Pd*, *Pt*, *Sc*, and *Ma*. On *K*, experimental N parents scored higher, but on all the rest, experimental A parents scored higher, in the more deviant direction.

Frequencies of MMPI Scales above T = 70

The next analysis involved computing frequencies of individual MMPI scores equal to or greater than a *T* score of 70, in order to study relative group differences in elevation of each scale by means of the chi-square test of proportions. This was done by counting the number of high points. A high point is here defined as any scale elevation greater or equal to *T* score = 70. For example, one profile (an experimental A mother) was coded 49 78' 10-3 26/5. This would be counted as four high points. Mothers and fathers were analyzed separately. None of the analyses for the mothers was significant. The proportions of high point codes *K* and *Pd* were significantly different for fathers, however. Experimental N fathers had more high point *K* ($p < .01$)

TABLE 3

RESULTS OF MMPI INDIVIDUAL SCALE ANOVAS,
EXPERIMENTAL A, EXPERIMENTAL N, AND
CONTROL MOTHERS AND FATHERS

Scale	Treatment (Rows)		Sex (Columns)
	A vs. N (df = 13/186)	C vs. A + N (df = 13/186)	(df = 13/186)
<i>L</i>	ns	ns	ns
<i>F</i>	ns	ns	ns
<i>K</i>	$F = 12.47^{***}$	ns	ns
<i>Hs</i>	$F = 12.14^{***}$	$F = 5.25^*$	$F = 5.84^*$
<i>D</i>	ns	$F = 10.39^{**}$	$F = 24.24^{****}$
<i>Hy</i>	ns	$F = 11.55^{**}$	$F = 23.38^{****}$
<i>Pd</i>	$F = 8.59^{**}$	$F = 16.77^{***}$	ns
<i>Mf</i>	ns	ns	$F = 472.8^{****}$
<i>Pa</i>	ns	ns	$F = 7.22^{**}$
<i>Pt</i>	$F = 7.36^{**}$	ns	$F = 7.71^{**}$
<i>Sc</i>	$F = 13.97^{***}$	ns	ns
<i>Ma</i>	$F = 9.67^{**}$	ns	$F = 6.73^*$
<i>Si</i>	ns	ns	ns

Note.—None of the interaction effects were significant.

* $p < .05$.
** $p < .01$.
*** $p < .001$.
**** $p < .00001$.

and Experimental A fathers had more high point *Pd* ($p < .001$).

Frequencies of "Low Mf" in Mothers

"Low *Mf*" has come to have a special meaning clinically for females, particularly in combination with high elevations in other scales. (It refers to a *high* score in the direction of femininity.) For this analysis, a female profile so classified was one in which *Mf* was the lowest clinical scale, as plotted on the standard *K*-corrected profile. The experimental A mothers had significantly more low *Mf* profiles than either the experimental N or the control mothers ($p < .05$).

Configural Differences

Configural differences in individual scores on *Hy* and *Pd* were analyzed by subtracting the *K*-corrected score on *Pd* from *Hy* for each parent. The *Pd* minus *Mf* difference score was obtained likewise. The distribution of both these differences appeared markedly skewed for fathers in all three groups (in the direction of greater differences), so median rather than mean values were calculated. The fathers and mothers were then compared separately using the Extension of the Median Test (Siegel, 1956).

Tests on both *Hy-Pd* and *Pd-Mf* differences were not significant for mothers. Fathers were significantly different on *Hy-Pd* at the .05 level, with the experimental A fathers higher on *Pd* than *Hy*, while both experimental N and control fathers were lower on *Pd* than *Hy*. The *Pd-Mf* differences were even greater (at the .001 level), with the experimental A and experimental N fathers having a higher *Pd* than *Mf* score, while the control fathers were reversed in direction, with *Mf* higher than *Pd*.

Summary of Results of MMPI Testing

To summarize results for fathers, experimental A fathers scored higher than either experimental N or control fathers on *Pd* and *Sc*. Experimental N fathers scored lower than either experimental A or control fathers on *Pt* but higher than either experimental A or control fathers on *K*. Experimental A fathers were higher on *Pd* than *Hy* using median difference scores, while experimental

N and control fathers were higher on *Hy* than *Pd*. Both experimental A and experimental N fathers were higher on *Pd* than *Mf* while control fathers were higher on *Mf* than *Pd*. The configural analyses confirmed the generally higher scores on *Pd* in the experimental A fathers, both absolutely and relatively to other scales.

Experimental A mothers scored lower than either experimental N or control mothers on *K*. They scored higher on Scales *Hs*, *Pd*, *Pt*, *Sc*, and *Ma* than either experimental N or control mothers. They had more low point *Mf* than either experimental N or control mothers. Both experimental A and experimental N mothers scored higher than controls on Scales *D* and *Hy*. The mothers in all three groups had much more similar group profiles than did the fathers.

DISCUSSION

Although the experimental A parents had the lowest socioeconomic status (SES) rating, they were not significantly different from the other two groups, as had been expected. This was contrary to the findings of the Philips and Zigler (1961) and Achenbach (1966) studies. The three groups were comparable on age of parents, age of children, and number of children, as well as SES.

The only demographic characteristic on which the clinic boys differed significantly was birth order. Neurotic boys tended to be first-borns, and aggressive boys, middle children. Sampson's (1965) review of the research in ordinal position presented very little comparable empirical data. Sampson concluded, however, that laboratory studies of aggression in younger children, such as those by Sears (1951), Haeberle (1958), and MacFarlane (1954), show that firstborn males are somewhat less aggressive than second-born males. The findings of the present study support these conclusions.

Previous research (Butcher, 1969) suggested that experimental parents combined would be more deviant than controls and that experimental A parents would be more deviant than experimental N parents. Both these expectations were fulfilled.

The greater deviance of the experimental A fathers, particularly on the *Pd* scale, is the

most striking finding in this study. In each statistical analysis, the *Pd* scale is markedly higher in this group, indicating the same kind of difficulty with control of overt aggression previously noted by Butcher (1966). The other notable scale elevation in the experimental A fathers was on *Sc*. Together with the high *Pd*, this indicates a group with poor impulse control and inability to tolerate meaningful close relationships, remaining cold, distant, hostile, and unpredictable.

The present investigation did not attempt in any way to directly examine parent and child identification; nevertheless, the symptoms in Achenbach's list of "aggressive" traits in boys such as temper tantrums, stealing, lying, and inadequate guilt feelings, appear very similar to the antisocial traits associated with elevation on the *Pd* scale.

The most notable finding in the experimental N fathers was the relatively high elevation on the *K* scale. Since the three groups were roughly comparable in background, this discrepancy cannot be attributed to differences in education or socioeconomic level. The experimental N fathers appear relatively more defensive as a group and less apt to admit shortcomings. This group may feel particularly unable to admit feelings of anxiety, or other "neurotic" symptoms. In other respects, the neurotic fathers appeared very similar to the control fathers.

Both experimental and control mothers followed the characteristic pattern of elevations on scales *Hy*, *Pd*, and *Pa* with relatively low *Mf*, previously found (e.g., Hanvik, 1959) to be characteristic of the typical mother of a child brought for counseling. The three groups differed relatively little from each other in configuration. The best explanation for this is that the mothers were not as important in influencing the development of characteristic pathology in their sons as were the fathers, which would agree with the usual conclusions of identity theory. The experimental A mothers did appear more deviant than either experimental N or control mothers, however, even though their group profile was similar. They were significantly higher on *Pd* and *Sc*, as were the experimental A fathers. It appears they also were hostile and had inadequate impulse control,

but their very low *Mf* suggests their hostility would be expressed by passive-aggressive maneuvers rather than directly.

The degree of pathology of experimental A parents is probably limited in this research by the requirement that parents be married and living together. As Loeb (1966) points out, parental separation and divorce is related to greater delinquent tendencies in children. (Children of divorced and separated parents had many more problems including aggression or acting-out behavior (64%) than children of continuously married parents (23%) ($p < .001$).) Although previous studies have found relatively greater pathology in mothers of "externalizers" or "acting-out" children (Cummings, 1967; Liverant, 1959; Rexford, 1966), this may well have resulted from the preponderance of mothers without husbands in these studies.

The most important clinical implication in this research is the necessity of including the father in assessment and treatment of boys, particularly aggressive boys. Also indicated is the importance of including fathers in further research in the area of parent-child personality interactions.

REFERENCES

- ACHENBACH, T. M. The classification of children's psychiatric symptoms: A factor analytic study. *Psychological Monographs*, 1966, 80(7, Whole No. 615).
- BANDURA, A., & WALTERS, R. H. *Adolescent aggression*. New York: Ronald Press, 1959.
- BUTCHER, J. N. MMPI characteristics of externalizing and internalizing boys and their parents. Paper presented at the First Conference on Recent Developments in the Use of the MMPI, Minneapolis, March 1966.
- BUTCHER, J. N. *MMPI: Research developments and clinical applications*. New York: McGraw-Hill, 1969.
- CUMMINGS, S. T., & CARSON, R. L. Maternal personality and the externalizing-internalizing symptom dimension in neurotic children. *Proceedings of the 75th Annual Convention of the American Psychological Association*, 1967, 2, 181-182.
- GLUECK, S., & GLUECK, E. *Unraveling juvenile delinquency*. New York: Commonwealth Fund, 1950.
- HAEBERLE, A. Interactions of sex, birth order, and dependency with behavior problems and symptoms in emotionally disturbed pre-school children. Paper presented at the meeting of the Eastern Psychological Association, Philadelphia, April 1958.

- HANVÍK, L. J., & BYRUM, M. MMPI profiles of child psychiatry patients. *Journal of Clinical Psychology*, 1959, 15, 427-431.
- LIVERANT, S. MMPI differences between parents of disturbed and nondisturbed children. *Journal of Consulting Psychology*, 1959, 23, 256-260.
- LOEB, J., & PRICE, J. R. Mother and child personality characteristics related to parental marital status in child guidance cases. *Journal of Consulting Psychology*, 1966, 30, 112-117.
- MACFARLANE, J. W., ALLEN, L., & HONZIK, M. P. A developmental study of the behavior problems of normal children between 21 months and 14 years. In *University of California publications in child development*. Vol. 2. Berkeley: University of California Press, 1954.
- PHILIPS, L., & ZIGLER, E. Social competence: The action-thought parameter and vicariousness in normal and pathological behaviors. *Journal of Abnormal and Social Psychology*, 1961, 63, 137-146.
- REXFORD, E. N. (Ed.) *A developmental approach to problems of acting out. Monographs of the Journal of the American Academy of Child Psychiatry*. Vol. 1. New York: International Universities Press, 1966.
- SAMPSON, E. E. The study of ordinal position: Antecedents and outcomes. In *Progress in experimental personality research*. Vol. 2. New York: Academic Press, 1965.
- SEARS, P. S. Doll play aggression in normal young children: Influence of sex, age, sibling status, father's absence. *Psychological Monographs*, 1951, 65(6, Whole No. 323).
- SIEGEL, S. *Nonparametric methods for the behavioral sciences*. New York: McGraw-Hill, 1956.

(Received September 10, 1968)

Manuscripts Accepted for Publication in the
Journal of Consulting and Clinical Psychology

- Effect of Nurses' Mode of Dress on Behavior of Psychiatric Patients Differing in Information Processing Complexity: H. Stephen Lef,*, Rudy V. Nydegger, and Mildred Buck: Harvard Medical School, 74 Fenwood Road, Boston, Massachusetts 02115.
- Etiquette of Verbal Conditioning: Bruce Denner*: Psychological Clinic, Indiana University, Bloomington, Indiana 47401.
- A Study of the Comparability of the WISC and the WAIS: Robert L. Simpson*: California State College at Fullerton, 800 North State College Boulevard, Fullerton, California 92631.
- Stereotypic Aspect in Masculinity-Femininity Measurement: Patricia W. Lunneborg*: Bureau of Testing, University of Washington, Seattle, Washington 98105.
- Relation between a Dimension of Internal-External Control and the MMPI with an Alcoholic Population: Allen Goss* and Thomas E. Morosko: Texas Research Institute of Mental Sciences, 1300 Moursund Avenue, Texas Medical Center, Houston, Texas 77025.
- Associative Interference in Schizophrenia as a Function of Paranoid Status and Premorbid Adjustment: Jay Y. Gonen*: School of Medicine and Dentistry, University of Rochester, 260 Crittenden Boulevard, Rochester, New York 14620.
- Quantified Judgments of Mental Health from the Rorschach, TAT, and Clinical Status Interview by Means of a Scaling Technique: Bertram P. Karon* and Paul O'Grady: Department of Psychology, Michigan State University, East Lansing, Michigan 48823.
- A Pilot Study on the Effectiveness of Systematic Desensitization with Neurologically Impaired Children with Phobic Disorders: Martin Ober and Robert F. Terwilliger*: Graduate Faculty of Political and Social Science, New School for Social Research, 66 West 12th Street, New York, New York.
- Effect of Differential Instructional Set on Sentence-Completion Responses: Floyd S. Irvin* and Marie L. Johnson: Institute for Psychosomatic Research and Training, 29th Street and Ellis Avenue, Chicago, Illinois.
- An Experimental Procedure for the Modification of Psychotic Behavior: Ivan Toby Rutner* and Charles Bugle: Psychological Service Center, University of Manitoba, Winnipeg, Canada.
- Culture Conflict in the Treatment of "Mental Illness" and the Central Role of Patient Leader: Alan F. Fontana* and Michel Corey: Veterans Administration Hospital, West Spring Street, West Haven, Connecticut.
- Mathematicians: The Creative Researcher and the Average PhD: Ravenna Helson* and Richard S. Crutchfield: Institute of Personality Assessment and Research, University of California, Berkeley, California 94720.
- Professional Research Commitment: A Symptom or a Syndrome? Richard L. Bednar* and Jeffrey G. Shapiro: Arkansas Rehabilitation Research and Training Center, University of Arkansas, Fayetteville, Arkansas 72703.
- Effects of Foreknowledge of Death in the Assessment from Case History Material of Intent to Die: Ralph W. Hood, Jr.*: Department of Psychology, South Dakota State University, Brookings, South Dakota 57006.

* Asterisk indicates for whom address is supplied.

LOCUS OF CONTROL IN PSYCHIATRIC PATIENTS

MARTIN HARROW¹ AND ANTHONY FERRANTE

Yale University

Acute psychiatric inpatients ($N = 128$) were administered the I-E scale during their first and seventh weeks of hospitalization. The overall sample's scores at Time Period 1 were within the average range found for nonpatient samples. Schizophrenics, however, were more external than nonschizophrenics ($p < .02$). Older patients were more internal ($p < .05$). I-E scores correlated significantly with scores on self-confidence and frustration. The results, with an acute sample, suggest that patients with greater psychopathology and fewer social skills (schizophrenics, younger patients, and to a slight extent, males) are more external. At Week 7, after clinical improvement, the overall sample's I-E scores had not changed significantly. Depressives, however, became more internal ($p < .02$). Schizophrenics differed from nonschizophrenics on change scores ($p < .001$). Female patients became more internal ($p < .01$).

An important dimension of personality which has been subject to increasing attention during the last 10 years is the variable called locus of control (Lefcourt, 1966; Rotter, 1966; Rotter, Seeman, & Liverant, 1962). Locus of control, a construct generated within Rotter's social learning theory, refers to the extent to which an individual perceives events in his life as being a consequence of his own actions, and thus under his personal control. It is assessed in terms of whether one believes that events in peoples' lives result from their own efforts, skills, and internal dispositions (internal control), or stem from external forces such as luck, chance, fate, or powerful others (external control).

This dimension of an individual's orientation toward the world has been investigated in many different normal populations and found to relate to social and behavioral factors (Gore & Rotter, 1963; Strickland, 1965). Externality is clearly associated with ethnic background and socioeconomic and educational levels and related to higher levels of conformity and passivity (Battle & Rotter, 1963; Crowne & Liverant, 1963; Graves, 1961). Conversely, internality has been thought to be related in curvilinear fashion to ego control (James, 1957).

Although there has been extensive interest in broadening knowledge about locus of con-

trol, little research has been conducted with psychiatric patients. It would seem important to determine the extent to which individuals with severe behavior disorders feel that events in their lives are a consequence of their own behavior. In addition, it would seem important to investigate whether various kinds of severely disturbed people (e.g., schizophrenics, depressives, character disorders) differ in the degree to which they perceive what happens to them as resulting from their own actions, and whether their attitudes in this area change during the course of treatment. One study which did use schizophrenic Ss found that their attitudes about personal control were external (Cromwell, Rosenthal, Shakow, & Zahn, 1961). This latter research, however, did not explore whether an external attitude is unique among schizophrenic patients or is a characteristic of all "mental" patients regardless of diagnosis. In addition, the relationship between locus of control in "mental" patients, and their age, sex, and level of chronicity were not explored, even though these latter factors may be important influences. A comprehensive investigation of the locus of control construct in psychiatric patients would be of interest to both the personality theoretician and the clinician.

The present research, part of a larger study of the personality profiles of psychiatric patients, presents a systematic investigation of attitudes about locus of control in a large sample of hospitalized patients. In this paper we are concerned with the following questions:

¹Requests for reprints should be sent to Martin Harrow, Department of Psychiatry, School of Medicine, Yale University, 333 Cedar Street, New Haven, Connecticut 06510.

General Characteristics of Locus of Control in Psychiatric Patients

1. What is the distribution of mental patients on the Internal/External continuum, and how does it relate to previously collected samples of "normals"?

2. Is there a relationship between these scores and diagnoses?

3. Are these scores related to age or sex?

Changes of Locus of Control Scores following 7 Weeks of Hospitalization

1. Do psychiatric patients become more internal as their manifest pathology decreases?

2. If there are any changes in internality or externality, are they related to the type of disorder of the patients (their formal diagnosis)?

3. Are changes in patients' attitudes about locus of control related to age or sex?

METHOD

Patient Population

The Ss were 128 consecutive admissions to a short-term, psychiatric inpatient service. The patients represented the major types of psychiatric disorders necessitating hospitalization (40 schizophrenics, 41 depressives, 34 character disorders, 5 manics, and 8 of varied diagnoses). Since the sample was later subdivided and examined separately according to diagnosis, special attention was devoted to this factor, with all patients being diagnosed according to standard psychiatric nomenclature. The diagnosis used for each patient was arrived at by consensus of two experienced clinicians (a board-certified psychiatrist and a clinical psychologist), each with over 6 years' experience in the field.

Demographically the patient sample was typical of the populations of many short-term, private psychiatric hospitals, with the major deviation from the usual population norms being a greater proportion of upper-middle-class and female patients. Thus, there were 45 males and 83 females among the patient population. The mean age of the sample was 31.2 years and the range was 12-72 years. The mean educational level of the sample was 14.0 years and the range was 7-20 years.

Setting

The research was carried out on the Acute Psychiatric Inpatient Service at Yale-New Haven Hospital. The unit relies heavily on milieu, group, and family therapy, with frequent use of ataractic medications (Astrachan, Harrow, Becker, Schwartz, & Miller, 1967; Harrow, Astrachan, Becker, Detre, & Schwartz 1967). The staff attempts to foster rapid social

relearning, control of symptoms, responsibility for behavior, and a more independent life adjustment after discharge (Norton, Detre, & Jarecki, 1963).

The overall orientation of the staff, and thus of the service, is an internal one, with emphasis on patients accepting responsibility. In this group-oriented ward, a high premium is placed on patients becoming active and responsible group participants, and members of the patient community. Within the framework of a therapeutic community, clinical improvement in a patient's condition is accompanied by increasing responsibility for oneself and for helping other disturbed patients. The primary therapeutic goal of the unit is the rapid return of the patient to the community with a reduction in symptomatology and the restoration of ability to function in the family, with peers, and at work (Detre, Kessler, & Jarecki, 1963).

Data Collection

The instrument for measuring locus of control was the Internal/External (I-E) scale (Rotter, 1966), which is a 29-item forced-choice scale with six buffer items. The 23 critical test items attempt to measure the degree that S perceives events in his life as being a consequence of his own actions, or beyond his personal control.

The questionnaire was administered to all patients during their first week of hospitalization.² Of the 128 patients who took the test during the first time period, 88 were still in the hospital 6 weeks later, at which point they were administered the test again for the second time.

The I-E test was one of a number of personality, attitude, and performance tests routinely administered to all patients at prescribed time periods over 8 weeks. In addition to the test material obtained from the patients, extensive daily and weekly ratings of all patients' clinical conditions and behavior were collected routinely from the patients' therapists and from other staff observers. The testing program, and the present study in particular, are part of an ongoing research project investigating the relationship between personality factors and behavior disorders.

RESULTS AND DISCUSSION

Initial Scores: Overall Sample

In regard to the data on the patient sample as a whole, the initial scores were expected to be external because of the following factors: (a) a reaction by the patients to acute psychiatric symptoms which infringe upon their sense of mastery, and thus may lead to exter-

² Several patients were discharged or transferred during the first week, before data could be collected from them. Several others, who were too disorganized to be tested at Week 1, were tested for the first time during the second to fourth week, and then for a second time 6 weeks later. Their data are included in the present sample.

TABLE 1
MEAN SCORES ON LOCUS OF CONTROL
AT TIME PERIOD 1

Group	<i>M</i>	<i>N</i>	<i>t</i>	
			Schizophrenics ^a	Manics ^b
Total sample	8.70	128	—	—
Schizophrenics	10.07	40	—	2.70*
Depressives	8.27	41	1.93	2.45*
Character disorders	8.06	34	1.91	1.96
Manics	4.20	5	2.70*	—
Others	9.50	8	.32	2.76*

^a The *t* scores represent comparisons between the schizophrenic subgroup and each of the other diagnostic groups listed.

^b The *t* scores represent comparisons between the manic subgroup and each of the other diagnostic groups listed.

* $p < .05$.

nality; (b) acute mental hospitalization which is usually at least partly coercive and thus not completely under the patients' own control; and (c) differences between these hospitalized patients and normal samples on long-standing factors such as personality difficulties and susceptibility to upsetting events, both of which would tend to give the patients a diminished sense of mastery and a diminished sense of control over their own fates.

The actual results presented in Table 1 show an initial mean locus of control score of 8.70 for the entire group of patients. This score lies within the range of mean scores presented by Rotter (1966) for several non-psychiatric samples. To check this further, the data from the current psychiatric patients were compared with those presented by Rotter (1966) from his sample of 1,180 Ohio State University students (mean I-E score of students = 8.29), and there were no significant differences ($t = 1.10$, $p > .20$). It must, of course, be kept in mind that the present patient sample is predominantly upper middle class with above-average intelligence (their mean IQ, based on the WAIS Information subtest, falls in the lower part of the superior range of intelligence), and samples of this type typically have been slightly more internal.

Despite these considerations, the data reported in Table 1 indicate that this group of hospitalized psychiatric patients does not dif-

fer significantly from nonpsychiatric samples. Thus, the results would seem to suggest that the above factors associated with psychiatric hospitalization do not influence locus of control in an external direction for the patient population, considered as a whole. It is still possible, however, that the above factors have an external influence, but that the effect is not uniform for the entire sample, and instead only influences subgroups of this population.

This possibility is investigated in the succeeding sections.

Initial Scores: Categorized According to Diagnosis

The data on locus of control for the various diagnostic groups also are presented in Table 1. The results fit a theoretically consistent pattern if one considers the premorbid history of many patients with these types of disorders and how their history (in terms of success or failure in dealing with environmental events) would be expected to influence the extent of externality.

The schizophrenic subgroup was the most external and the mean score of these patients (10.07) was significantly more external than the mean (8.07) for the total sample of non-schizophrenic patients ($t = 2.51$, $df = 126$, $p < .05$). Similar results (in terms of schizophrenics being external) were obtained by Cromwell et al. (1961), although their schizophrenic sample was not compared to a sample of nonschizophrenic psychiatric patients. Looked at in terms of their history, the schizophrenic sample should contain a greater percentage of patients who have had long-standing difficulties in adjustment and have experienced chronic disappointment. Furthermore, this disorder includes symptoms such as delusions and paranoid ideation which are threatening to others, produce greater social stigma, and may lead to a sense of personal futility, with some resulting externality.

The depressives are a group with a presumably better premorbid adjustment, and a less-chronic, long-term disorder than the schizophrenics. Although their current depressive illness (with symptoms such as worthlessness and self-depreciation) may have influenced the scores in the external direction, their

scores tend toward the internal as compared to schizophrenics ($p < .10$, see Table 1).

The patients with character disorders may have some long-standing personality difficulties, but these are generally ego syntonic, and many of these patients have performed in a competent manner in the social and occupational areas over many years. A large number of these patients have had difficulty in only one or two select areas, often related to their life styles. Despite their admission to the hospital, arising from some crisis which they were presumably unable to handle, these patients still tended to be more internal than the schizophrenics ($p < .10$, Table 1).

At the height of their disorder (which is frequently at the time of hospitalization), manic patients display grandiose thoughts regarding their own ability to successfully deal with life. Their extremely internal orientation ($M = 4.20$) is consistent with this temporary confidence in their own capacity to meet life's challenges. The manic patients were significantly more internal than the total sample of nonmanic patients ($t = 2.44$, $df = 126$, $p < .05$).

The present data were obtained with largely acute, middle-class patients. It is quite possible that a sample of more chronic, less middle-class patients might have produced different (perhaps more external) results.

Initial Scores: Age and Sex

Table 2 presents the results when the total sample was categorized according to age and sex. Since the data on age and sex differed in several cases according to whether or not the patients were schizophrenic, the results for the schizophrenic and nonschizophrenic samples are presented separately in the table. When the data were analyzed according to age and sex, it was found that the younger patients were significantly more external ($p < .05$) and there was a nonsignificant trend for males to be more external than females. Although these tendencies held for the patient population as a whole, there were some complex interactions between age, sex, and diagnosis.

Thus, while older patients were significantly more internal, this was largely due to the overwhelming trend toward greater internality

TABLE 2

MEAN INITIAL SCORES ON LOCUS OF CONTROL
ACCORDING TO AGE AND SEX

Patients	Schizophrenics		Nonschizophrenics	
	<i>M</i>	<i>N</i>	<i>M</i>	<i>N</i>
Younger*	9.75	24	9.30	39
Older*	10.13	16	7.02	49
Male	10.77	22	8.17	24
Female	8.83	18	7.98	64

Note.—Young patients versus old patients: $t = 2.24$, $p < .05$.
Young nonschizophrenics versus old nonschizophrenics: $t = 2.71$, $p < .01$. Males versus females: $t = 1.56$.

* The sample was divided arbitrarily into younger patients (24 years and younger) and older patients (25 years and older).

among the older nonschizophrenic population as contrasted to the younger nonschizophrenic population. In many ways the scores for this older nonschizophrenic population of patients were most similar (generally internal) to what might be expected of a nondisturbed upper-middle-class population. This is not surprising since the older nonschizophrenic patients were generally less "sick" (via independent ratings of psychopathology by their psychiatrists and by nurses) than the younger or schizophrenic patients. In contrast, the younger patients, in addition to their greater pathology, frequently have fewer social skills. Perhaps of key importance, most of the older nonschizophrenic patients have a greater backlog of successful experiences. Thus, in a way, this subportion of the total hospital population contains the patients most similar to "normals" in terms of previous competence and achievement.

Despite the overall results of significantly greater externality among younger patients, this age-related trend was not found among the schizophrenic population. The older schizophrenics, who represented a more chronic population, were rather external. Most of these older schizophrenics have had long-standing difficulties in adjustment. It may be that among the older schizophrenics, with increasing age a backlog of adjustment difficulties and social disappointments have led to a sense of personal futility and have contributed to their more external stance.

It should be noted that the results on age, sex, and diagnosis fit a general pattern in

which the less competent patients with greater psychopathology and fewer social skills (schizophrenics, younger patients, and, to a slight extent, males) were the most external. This would generally fit a formulation in which general internal and social competence, and a backlog of successful experiences, would be one set of factors leading to increased internality. Evidence to support this formulation was found when the present patient samples' initial I-E scores were correlated with their scores on a scale for self-confidence used previously by Brim, Glass, and Lavin (1962), and one related to frustration over previous failures used by Srole (1962). The correlation between the I-E scores and self-confidence scores was .23 ($df = 121$, $p < .02$), and that between the I-E scores and frustration scores was .30 ($df = 121$, $p < .01$).

Change Scores: Overall Sample

It was expected that at Time Period 2 the patients would have changed in the direction of decreased externality. The above formulation is based on the hypothesis that as symptomatic improvement occurs, patients will become capable of greater mastery and feel more capable and adequate (Harrow, Fox, Markhus, Stillman, & Hallowell, 1968). This, in turn, could lead to increased internality. It is also possible that at a university teaching hospital, where the patients and staff have frequent regular contact, patients will identify to some degree with the professional staff's more internal values. Other evidence (Astrachan, Harrow, & Flynn, 1968; Astrachan, Schwartz, Becker, & Harrow, 1967;

Harrow, Astrachan, Becker, Miller, & Schwartz, 1967) indicates some acceptance of staff attitudes and values by patients in such a hospital setting.

The mean change scores for patients administered the I-E test twice, and the tests of significance associated with these changes are presented in Table 3. There was a slight, nonsignificant, trend for the patient sample as a whole to change in the direction of increased internality during the 6-week period. The surprisingly small (nonsignificant) change, coupled with the relatively large size of the sample, lead to the rejection of the hypothesis of increased internality for the patient sample considered as a whole. These results are all the more surprising since recent data of Hersch and Scheibe (1967) suggest a trend toward increased internality on test-retest over time among a "normal" sample. This trend (probably related, in part, to the test instrument) should be kept in mind in considering the subsequent data presented on the patients' change scores.

Change Scores: Categorized according to Diagnosis

When the patient sample is divided into the various diagnostic subgroups and the change scores analyzed separately for each group (Table 3), marked differences emerge.

Contrary to the general trend toward internality, the schizophrenics (whose scores at Time Period 1 were the most external) showed a nonsignificant trend toward increased externality. The change scores for the schizophrenic subgroup differed significantly from those of the rest of the patient population ($t = 4.14$, $df = 86$, $p < .001$). The absence of a change toward internality is not surprising when considered in conjunction with several features related to the schizophrenics' clinical status and the outlook that their pathology could engender. In terms of the degree of pathology in this group at the second testing period, many of these patients would still be labeled as very "sick." Thus, due to their clinical condition one might not expect to see the sense of increased mastery, potency, and subsequent increased internality which occurred in some other diagnostic groups.

TABLE 3
CHANGE SCORES ON LOCUS OF CONTROL

Group	Mean change	N	t^a
Total sample	-.44	88	-1.38
Schizophrenics	.72	29	1.65
Depressives	-1.12	25	-2.71*
Character disorders	-1.22	23	-1.64
Manics	3.00	5	1.89
Others	-3.17	6	-4.53**

^a The t scores represent differences between the obtained mean and zero (no change).

* $p < .02$.

** $p < .01$.

In addition, despite some clinical improvement, the schizophrenics have less to look forward to. There is usually more pessimism concerning the prognosis of these patients than is the case with other diagnostic groups. This can be seen in the greater difficulty in arranging for the posthospital psychotherapy of schizophrenics, due to the greater reluctance of many outpatient therapists to treat them.

In contrast to the schizophrenics, the depressive group showed a strong trend toward increased internality during the 6-week period ($p < .05$). These results indicate that as their depressive symptoms diminish, these patients become more internal. It is not surprising that the return to their premorbid levels of self-esteem and control are reflected in increased confidence in their ability to master the environment, and a resulting change toward increased internality.

The character disorders also tended to become slightly more internal during the 6-week period. It would have been difficult to make a precise prediction concerning the change scores for this group since they represent a heterogeneous set of character types and a mixed group of problems. On a post hoc basis it might be noted that while much of their character pathology remains after hospital treatment, at least the acute reason for their hospitalization has been relieved and they are in the process of returning to their previous, moderately successful, adaptation.

In general, the nonschizophrenic patients (other than manic patients) tended to become more internal as they improved during the 6-week period between testings. These results fit the overall picture of increased internality with an increasing sense of competence (or decreasing incompetence), and this is in accord with the earlier hypothesis.

The tendency for the manic patients to become more external (Table 3) is in accord with clinicians' observations of these patients. During hospitalization they recover from their original grandiosity concerning their plans, capabilities, and futures, and become more cognizant of their limitations. The external shift toward more "normal" locus of control is fully in accord with this increased realism.

TABLE 4

CHANGE SCORES ON LOCUS OF CONTROL
ACCORDING TO AGE AND SEX

Patients	Schizophrenics			Nonschizophrenics		
	M	N	t^b	M	N	t^b
Youngers	.83	18	1.57	-1.00	27	-1.52
Older ^a	.36	11	.46	-1.00	32	-1.98
Male	1.29	17	2.48*	-.06	17	-.07
Female	-.25	12	-.35	-1.38	42	-3.28**

Note.—Young patients versus old patients: $t = .61$. Male patients versus female patients: $t = 2.80$, $p < .01$. Young males versus young females: $t = 3.30$, $p < .01$.

^a The sample was divided arbitrarily into younger patients (24 years and younger) and older patients (25 years and older).

^b The t scores represent differences between the obtained mean and zero (no change).

* $p < .05$.

** $p < .01$.

Change Scores: Age and Sex

The patient sample also was categorized according to age and sex to determine whether there were any relationships between these factors and the extent of change in locus of control during the 6-week period. The results again are presented separately for the schizophrenic and nonschizophrenic patients. There were no significant differences between younger and older patients on the change scores.

The results on sex (Table 4) showed a general trend for a change toward internality among the females ($t = 3.09$, $df = 53$, $p < .01$) with a mixed picture among the males, who tended to become slightly more external. The difference between the male and female change scores was significant ($t = 2.86$, $df = 86$, $p < .01$). The greatest sex difference occurred when the group of patients under 25 years of age were compared according to sex ($t = 3.32$, $df = 43$, $p < .01$). The younger females showed a strong trend toward internality ($p < .01$), while the younger males tended to become somewhat more external ($p < .10$). Despite the young males' clinical improvement during the 6-week period between testings, many were preparing to face again a society where much is demanded of them. Thus, many upper-middle-class males feel strong pressures to achieve (due to the expectations of friends and relatives, and to their own aspirations). These pressures, coupled with their own uncertainty over whether they are able to meet the pressures successfully, could easily arouse potential feelings of frustration and powerlessness.

The younger females also experience similar but less intense pressures. Many do not feel as much social pressure to take the initiative and succeed. Thus, a number may have felt more competent than the males to meet the social challenges they will face on discharge, with resulting increased internality.

Test-Retest Reliability

It should be noted that while the data for a number of subgroups showed some change over time, the patients' scores at Time Period 1 (within a week after acute hospitalization) were closely related to their scores at Time Period 2 (6 weeks later, when the symptomatology had been drastically reduced). In terms of reliability, the test-retest correlation was .75 ($df = 86$) for the current sample of "mental" patients. This compares favorably with results obtained with more "normal" college students. Thus the test-retest correlations were .55 and .72 for two samples of students presented by Rotter (1966), and the test-retest coefficients ranged from .47 to .84 for 10 groups of college students tested by Hersch and Scheibe (1967).

Further subdivision of the sample into discrete types of patients showed similar results. The test-retest correlation for the nonschizophrenic patients was .67 ($df = 58$), and that for the schizophrenic patients was .87 ($df = 26$). Similarly the test-retest correlation for the older patients was .75 ($df = 41$), and for the younger patients was .73 ($df = 43$). Finally, the test-retest correlation for male patients was .79 ($df = 32$), and that for female patients was .74 ($df = 52$).

These high test-retest reliabilities and the comparative closeness of the scores at the two time periods indicate that while acute psychiatric hospitalization and acute symptomatology influenced some of the scores in a more external or internal direction, other, perhaps more basic, characteristics were major determinants of the scores for most of the patients.

REFERENCES

- ASTRACHAN, B. M., HARROW, M., BECKER, R. E., SCHWARTZ, A. H., & MILLER, J. C. The unled group as a therapeutic tool. *International Journal of Group Psychotherapy*, 1967, 17, 178-191.
- ASTRACHAN, B. M., HARROW, M., & FLYNN, H. R. Influence of the value system of a psychiatric setting on behavior in group therapy meetings. *Social Psychiatry*, 1968, 3, 165-172.
- ASTRACHAN, B. M., SCHWARTZ, A. H., BECKER, R., & HARROW, M. The psychiatrists effect on the behavior and interaction of therapy groups. *American Journal of Psychiatry*, 1967, 123, 1379-1387.
- BATTLE, E., & ROTTER, J. B. Children's feelings of personal control as related to social class and ethnic group. *Journal of Personality*, 1963, 31, 482-490.
- BRIM, O., GLASS, D., & LAVIN, D. *Personality and decision process: Studies in the social psychology of thinking*. Stanford: Stanford University Press, 1962.
- CROMWELL, R., ROSENTHAL, D., SHAKOW, D., & ZAHN, T. Reaction time, locus of control, choice behavior and descriptions of parental behavior in schizophrenic and normal subjects. *Journal of Personality*, 1961, 29, 363-380.
- CROWNE, D. P., & LIVERANT, S. Conformity under varying conditions of personal commitment. *Journal of Abnormal and Social Psychology*, 1963, 66, 547-555.
- DETRE, T. P., KESSLER, D. R., & JARECKI, H. G. The role of the general hospital in modern community psychiatry. *American Journal of Orthopsychiatry*, 1963, 33, 690-700.
- GORE, P. M., & ROTTER, J. B. A personality correlate of social action. *Journal of Personality*, 1963, 31, 58-64.
- GRAVES, T. D. *Time perspective and deferred gratification pattern in a tri-ethnic community*. (Research Rep. No. 5, Tri-Ethnic Research Project) Boulder: University of Colorado, Institute of Behavioral Science, 1961.
- HARROW, M., ASTRACHAN, B. M., BECKER, R. E., DETRE, T., & SCHWARTZ, A. H. An investigation into the nature of the patient-family therapy group. *American Journal of Orthopsychiatry*, 1967, 37, 888-899.
- HARROW, M., ASTRACHAN, B. M., BECKER, R. E., MILLER, J. C., & SCHWARTZ, A. H. Influence of the psychotherapist on the emotional climate in group therapy. *Human Relations*, 1967, 20, 49-64.
- HARROW, M., FOX, D. A., MARKHUS, K. L., STILLMAN, R., & HALLOWELL, C. B. Changes in adolescents' self concepts and their parents perceptions during psychiatric hospitalization. *Journal of Nervous and Mental Disease*, 1968, 147, 252-259.
- HERSCH, P. D., & SCHEIBE, K. E. Reliability and validity of internal-external control as a personality dimension. *Journal of Consulting Psychology*, 1967, 31, 609-613.
- JAMES, W. H. Internal versus external control of reinforcement as a basic variable in learning theory. Unpublished doctoral dissertation, Ohio State University, 1957.
- LEFCOURT, H. M. Internal versus external control of reinforcement: A review. *Psychological Bulletin*, 1966, 65, 206-220.

- NORTON, N. M., DETRE, T. D., & JARECKI, H. G. Psychiatric services in general hospitals: A family-oriented redefinitive. *Journal of Nervous and Mental Disease*, 1963, 136, 475-484.
- ROTTER, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 1966, 80(1, Whole No. 609).
- ROTTER, J. B., SEEMAN, M., & LIVERANT, S. Internal versus external control of reinforcements: A major variable in behavior therapy. In N. F. Washburne (Ed.), *Decisions, values, groups*. Vol. 2. London: Pergamon Press, 1962.
- SROLE, L. (Ed.) *Mental health in the metropolis*. New York: McGraw-Hill, 1962.
- STRICKLAND, B. The prediction of social action from a dimension of internal-external control. *Journal of Social Psychology*, 1965, 66, 353-358.

(Received September 13, 1968)

(Continued from page 581)

- Responses of A and B Subjects to Normal, Neurotic, Schizophrenic, and Ambiguous Communications: Robert J. Hoffnang and Leonard S. Stein*: Division of Psychology, University of Rochester School of Medicine, Rochester, New York 14620.
- Comparing Behavior Modification Approaches to Habit Decrement—Smoking: M. K. Wagner* and R. A. Bragg: Veterans Administration Hospital, Salisbury, North Carolina 28144.
- Creativity, Birth Order, and Preference for Symmetry: Russell Eisenman* and Nancy H. Robinson: Department of Psychology, Temple University, Philadelphia, Pennsylvania 19122.
- Private Practice Research: Wirt M. Wolff*: Division of Psychology, Southwestern Medical School at Dallas, 5323 Harry Hines Boulevard, Dallas, Texas 75235.
- Types of Stressful Situations and their Relation to Trait Anxiety and Sex: William F. Hodges* and James P. Felling: Department of Psychology University of Colorado, Boulder, Colorado 80302.
- Effects of Religious Affiliation, Religious Involvement, and Sex on the Social Desirability Ratings of MMPI Religion Items: Malcolm D. Gynther,* Bernadette G. Gray, and Milton E. Strauss: Department of Psychology, Saint Louis University, 221 North Grand Boulevard, Saint Louis, Missouri 63103.
- Ego Functioning and Memory Style: Lester Alston*: Department of Psychology, Hunter College, 695 Park Avenue, New York, New York.
- Effect of Three Conditions of Administration on "Trait" and "State" Measures of Anxiety: George J. Allen*: Department of Psychology, University of Illinois, Urbana, Illinois.
- Relation of the A-B Distinction and Trust-Distrust Sets to Addict Patients' Self-Disclosures in Brief Interviews: Juris I. Berzins,* Wesley F. Ross, and Daniel I. Cohen: Department of Psychology, University of Kentucky, Lexington, Kentucky 40506.
- Preliminary Validation Evidence for the Group Therapy Questionnaire: Daniel B. Wile,* Gary D. Bron, and Herbert B. Pollack: Counseling Services, California State College, 25800 Hillary Street, Hayward, California 94542.
- An Initial Study Using Fixed-Role and Rational-Emotive Therapy in Treating Public-Speaking Anxiety: Thomas O. Karst* and Larry D. Trexler: Department of Psychology, Temple University, Philadelphia, Pennsylvania 19122.
- The Columbia Mental Maturity Scale as a Test of Concept Formation: Jeanette Reuter* and Joanne Hyman: Psychology Department, Kent State University, Kent, Ohio 44240.
- Institutional and Community Adjustment of Delinquents following Counseling: Richard C. Sowles* and John H. Gill: Utah State Industrial School, Box 41, Ogden, Utah.
- Relationship between Patient-Therapist Compatibility and Treatment Effectiveness: Suzanne Marie Gassner*: Psychology Department, University of Massachusetts, 100 Arlington Street, Boston, Massachusetts 02116.
- The Long Search for a Short WAIS: Stop Looking: Margaret Barron Luszki,* William Schultz, H. Robert Laywell, and Robyn N. Dawes: Veterans Administration Hospital, 109 Bee Street, Charleston, South Carolina 29403.
- Predisposition to Fantasy and Situational Variations as Determinants of Attention in Reaction Time Tasks: Herbert M. Lefcourt* and Jerome M. Siegel: Department of Psychology, University of Waterloo, Waterloo, Ontario, Canada.
- Relationship of Reported Memories of Early Experiences with Parents on Interview Behavior: J. R. Baugh,* G. R. Pascal, and T. B. Cottrell: Suite 403 Medical Tower, 440 East Woodrow Wilson, Jackson, Mississippi 39216.
- Prediction of Rehabilitation Outcomes from Psychometric Parameters in Left Hemiplegics: Yehuda Ben-Yishay,* Louis Gerstman, Leonard Diller, and Albert Haas: Institute of Rehabilitation Medicine, 400 East 34th Street, New York, New York 10016.

* Asterisk indicates for whom address is supplied.

(Continued on page 620)

STUDY OF PSYCHOLOGICAL CONTENT IN THE COMMUNICATIONS OF SUICIDAL INDIVIDUALS¹

ALLEN R. DARBONNE²

Central Research Unit, Veterans Administration Center, Los Angeles

A test was made of the general hypothesis that the psychological content of communications of individuals who commit suicide differs from that of non-suicidal normal individuals. A comparison was also made of suicide-threatening outpatients with suicide committers and nonsuicidal normals. The message form selected for analysis was the suicide note. The Ss were American-born, white, adult males matched for age and occupational level. Overall results were in support of the major hypothesis. The content of communications of suicide committers reflected heightened dependency needs, difficulty maintaining interpersonal relationships, a high degree of activity, and an expressive style of veiled aggression which includes hurting the self in order to hurt and manipulate others. Suicide threateners were predictably in some ways like suicide committers and in certain other ways more like nonsuicidal normals.

Every 30 minutes, a person in the United States kills himself. The shock wave of these events, probably more than with the experience of any other human event, is consistently followed by the serious question, "Why?" The first attempt to deal with this question usually involves pointing to a precipitating stress just preceding the suicide. But students of suicide are becoming increasingly aware that the dynamics of suicide cannot be understood in terms of precipitating stress alone, and that there may be certain features of the suicide's personality which characterize his self-destructive approach to life and to other human beings (Farberow & Shneidman, 1961).

The primary purpose of the present study is to test the general hypothesis that the psychological content of communications of individuals who commit suicide differs from that of nonsuicidal individuals. This hypothesis is based on the assumption that the psychological content of communications of an individual is relatively stable and reflects his life style. This view has been expressed by several theorists in recent years, including

Ruesch (1961), Bateson, Jackson, Haley, and Weakland (1956), and Sullivan (1953).

The practical implication for suicide research from a communications point of view is that it reduces the distance usually found between theory and observations. Since a therapist's primary observations are of verbal and nonverbal communications, and his primary treatment method involves communication, it seems reasonable that theory used by therapists should be related to these same operations.

The second goal of the present study is to compare the content in communications of individuals who have threatened suicide with that of individuals who have actually committed suicide and with that of nonsuicidal individuals.

Adler's (1958) theory of the life style of suicidal individuals was used as a framework from which 19 hypotheses (presented in the Results section) concerning the communications content and style of suicidal individuals were derived. Adler's theory was selected as the framework because of its comprehensive quality and its explicit statements regarding life style which could be easily rephrased in operational terms concerning communications style and content.

These 19 hypotheses concerned predictions that the content of communications of suicide committers would reflect four general psychological characteristics: (a) heightened depen-

¹ This article is based on a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy under the chairmanship of Georgene Seward in the Department of Psychology at the University of Southern California.

² Also at Reiss-Davis Child Study Center, Los Angeles. Requests for reprints should be sent to Allen R. Darbonne, Veterans Administration Center, Central Research Unit, Building 13, Wilshire and Sawtelle Boulevards, Los Angeles, California 90073.

dency needs, (b) low self-esteem and self-centered goals, (c) a high degree of activity, and (d) an expressive style of veiled aggression which includes hurting the self in order to hurt and manipulate others.

Thus, it was predicted that the messages of suicide committers would show more evidence of these variables than those of nonsuicidal normals. Further, it was expected that persons who threatened suicide would communicate in certain specific ways similar to suicide committers and in other specific ways similar to nonsuicidal normals. In other words the prediction here was that the suicide threat group would always fall between the commit-suicide and the nonsuicidal groups relative to the observed frequency of each communications content item.

METHOD

The message form which was selected to be analyzed and compared was the suicide note. This was done because of the suicide note's unique quality of being a message which is primarily structured by the writer and is available from a large number of persons who have taken their lives. Studies analyzing the content of suicide notes have demonstrated the validity and reliability of the method. Shneidman and Farberow (1957b), who introduced the technique of suicide note analysis, reported on a comparison between genuine and simulated suicide notes in terms of Mowrer's concepts of discomfort and relief. They found the discomfort-relief quotient (DRQ) interesting, but of limited value in differentiating suicidal and nonsuicidal individuals. This was primarily attributed to the DRQ's insensitivity to the quality of discomfort and relief. They also reported that data analysis has shown no significant difference in economic, social-cultural, and personal factors between suicide committers who leave notes and those who do not.

In another study, Shneidman and Farberow (1957c) again compared the content of suicide notes and pseudosuiticide notes. They concluded that suicidal individuals use a type of logic which involves a fallacious identification. They term it a "psycho-semantic fallacy." This fallacious identification between the self as experienced by the self (I_s) and the self as it feels itself experienced by others (I_o) results in the suicide's acceptance of erroneous premises and invalid conclusions. One of these is the assumption that the suicide committer will be able to experience after death the effect his expiration will have on others. In a study concerning suicide and age, Farberow and Shneidman (1957) rated actual suicide notes on the basis of Menninger's (1938) concept of the (a) wish to kill, (b) wish to be killed, and (c) the wish to die. They found more intense interpersonal motives expressed in the notes

of younger suicides and more chronic depressive feelings expressed by older suicides. In an analysis of 948 genuine suicide notes in Los Angeles County, Shneidman and Farberow (1960) found reasons for suicides, and affect expressed in notes differed significantly between nine sociological area types.

Darbonne (1969) studied the age variable in relation to content of actual suicide notes and found significant differences between age groups. Osgood and Walker (1959) studied the effect of motivation level on language behavior, assuming actual suicidal individuals are under heightened motivation. They compared actual suicide notes, simulated suicide notes, and ordinary letters to friends and relatives. Tuckman, Kleiner, and Lavell (1959) investigated the emotional content of actual suicide notes and found that 50% of the notes contained positive affect, 25% expressed negative affect, and 25% were neutral in affect. All investigators using suicide notes have reported high reliability with this technique. In a validation study, Tuckman, Kleiner, and Lavell (1960) analyzed 63 suicide notes and found reasons for suicide stated in the notes in agreement with information of an informant in over 90% of the cases.

The Ss for the study were (a) 40 individuals who had actually committed suicide (C/S group), (b) 40 individuals who indicated they were not then and had never been suicidal (N/S group), and (c) 31 psychotherapy outpatients who were threatening to commit suicide (T/S group). All Ss were American-born, white males. In addition, the C/S group and the N/S normal group were matched for age (within 5 years) and for occupational level.

The actual suicide notes were randomly selected (from the files of the Los Angeles County coroner) and included American-born, white males between the ages of 20 and 59 who had committed suicide between July 1, 1963, and June 30, 1964. Suicide notes were obtained from the nonsuicidal (N/S) and the suicide threateners (T/S) by asking them to write the suicide note they would write if they were going to take their own lives. Extensive control procedures were used to avoid the possibility of individual adverse posttask reaction to writing a suicide note. A standard form was used to analyze the suicide notes in terms of the 19 hypotheses. A method of analysis was used which controlled for length of the suicide note. All analyses were done blind, each note having been coded, retyped on similar typewriter and paper, and shuffled prior to being seen by the scorer. Interrater reliability tested by the percentage of agreement method yielded a mean agreement on the 19 categories of 87.9%.

RESULTS

The overall results of the comparison between the suicide-commit and nonsuicidal groups are given in Table 1. Ten of the 19 hypotheses were supported by the data ($p < .05$), a number greater than would be ex-

TABLE 1
RESULTS OF COMPARISON OF C/S AND N/S GROUPS

Hypotheses	Frequency C/S	Frequency N/S	χ^2	p
1a. Need others do for him	29	9	18.10	.0005
1b. Needs emotional support	18	8	4.62	.025
1c. Loss of something	12	6	1.79 <i>ns</i>	(.10)
1d. Parent figures important	18	10	2.69	.05
2a. Low self-esteem	12	11	.00	<i>ns</i>
2b. At mercy of external forces	15	14	.00	<i>ns</i>
2c. Concern about others	25	11	"	.01 ^b
2d. Self-referents	18	22	"	<i>ns</i> ^b
2e. Break in social ties	11	4	2.95	.05
3a. Communicate more	23	17	"	<i>ns</i> ^b
3b. Motor-action oriented	24	12	"	.05 ^b
3c. High activity—suicide method	15	14	.00	<i>ns</i>
3d. Activity concern > contemplation	19	3	14.11	.0005
4a. Others hurt by past behavior	6	3	.50	<i>ns</i>
4b. Concern regarding blame for suicide	12	2	7.01	.005
4c. Suffering mentioned	16	19	.20	<i>ns</i>
4d. Targets named	35	18	14.31	.0005
4e. Express direct anger	2	3	.00	<i>ns</i>
4f. Veiled hostility	12	4	3.83	.05

^a Compared by Mood's (1950) Median Test.

^b Level of significance evaluated by reference to Mainland and Murray (1952) table.

pected to occur by chance. The data also showed a trend in the direction of support for Hypothesis 1c ($p < .10$). Of all 19 hypotheses, only two cases were found in which the observed frequencies were in the opposite direction of that which was expected. These were Hypotheses 2d and 4c.

The results of the suicide-commit and suicide-threat group comparisons and the non-suicidal and suicide-threat group comparisons

in terms of those variables which differentiated the suicide-commit and nonsuicidal groups are given in Table 2. The overall results of this comparison support the hypothesis that suicide threateners, as a group, are in some ways like suicide committers, but in other ways unlike them, and that these similarities and differences are detectable in the psychological content of their communications.

Since there were 19 minor hypotheses, in

TABLE 2
RESULTS OF COMPARISON OF C/S AND T/S GROUPS AND N/S AND T/S GROUPS

Hypotheses	Comparing C/S with S/T		Comparing N/S with S/T	
	χ^2	p	χ^2	p
1a. Need others do for him	16.14	.0005	.12	<i>ns</i>
1b. Needs emotional support	.00	<i>ns</i>	3.15	.05
1c. Loss of something	1.16	(.15)	.00	<i>ns</i>
1d. Parent figures important	.00	<i>ns</i>	1.58	<i>ns</i> (.15)
2c. Concern about others	1.13	<i>ns</i>	3.28 ^a	.05
2e. Break in social ties	.21	<i>ns</i>	3.45	.05
3b. Motor-action oriented	4.34	.025	.42 ^a	<i>ns</i>
3d. Activity concern > contemplation	3.67	.05	2.15	<i>ns</i> (.10)
4b. Concern regarding blame for suicide	.19	<i>ns</i>	3.42	.05
4d. Targets named	8.00	.005	.34	<i>ns</i>
4f. Veiled hostility	.00	<i>ns</i>	4.15	.025

^a Compared by Mood's (1950) Median Test.

consideration of brevity and clarity, rather than stating each hypothesis and then coming back and separately giving the findings, the findings are reported (in parenthesis) after each hypothesis is presented.

1. Pampered Life Style (Dependency)

1a. The first hypothesis of the group related to pampered life style was that the messages of suicidal individuals were expected to reflect the writers' need to have things done for them. Content related to others doing something for the communicator or indicating the communicator expects or demands that others fulfill his wishes will be found more frequently in the suicide notes of the C/S group than the N/S group. (This hypothesis was supported by the data. In terms of this item, the T/S group was similar to the N/S group and differed statistically from the C/S group.)

1b. The next hypothesis was that the messages of suicidal individuals would reflect their concern regarding the emotional support of others. Content related to whether or not the communicator feels others have supported him emotionally or have let him down will be found more frequently in the suicide notes of suicidal individuals than those of N/S individuals. (This hypothesis was also supported by the data. The T/S group was more like the C/S group than the N/S group relative to this item.)

1c. A third hypothesis was that the suicidal person's difficulty adapting to loss will be evident in his messages. Mention of having lost something and being unable to do without it will be found more frequently in the suicide notes of suicidal individuals than N/S individuals. (Support was found for this hypothesis at the .10 level of significance. The T/S group was similar to the N/S group and showed a trend toward differing [$p = .15$] from the C/S group regarding this item.)

1d. The fourth hypothesis concerning dependency was that the heightened importance of parent figures will be notable in communications of suicidal persons. References to parents and/or authority figures will be found more frequently in the suicide notes of suicidal individuals than in those of N/S individuals. (Again, this prediction was supported

by the data. The T/S group was similar to the C/S group and showed a trend toward differing [$p = .15$] from the N/S group in terms of this hypothesis.)

2. Inferiority Feelings and Self-Centered Goals

2a. It was hypothesized that content indicating low self-esteem would be found more frequently in the messages of suicidal individuals than those of nonsuicidal individuals. (This hypothesis was not supported by the data. Content reflecting low self-esteem was very prevalent in the C/S notes, but also in the notes of all other groups.)

2b. It was expected that the notes of the suicide committers would, more frequently than those of nonsuicidals, reflect a feeling of not being in control of their own destiny and a feeling of being at the mercy of forces outside of themselves. (This expectation was not supported by the data.)

2c. It was posited that the suicidal individual's dependence on others for need satisfaction would result in his referring to other persons in the suicide note more frequently than would be found in the messages of nonsuicidal individuals. (This expectation was supported by the data. The T/S group was like the C/S group and differed from the N/S group in this respect.)

2d. It was predicted that the suicidal individual's self-centered thinking would result in his using self-referents more frequently than would nonsuicidal individuals. (This prediction was not supported by the data.)

2e. Supported by the data was the hypothesis that the suicide's self-centered goal would interfere with the formation of lasting social ties. (Notes of the C/S group more frequently indicated a loss or absence of social ties or a break in interpersonal relations than did the N/S notes. Again the T/S group was more like the C/S group than the N/S group.)

3. Degree of Activity

It was felt this concept was related to the "tendency to act-out" notion. This is a crucial concept whenever there is an attempt to differentiate persons who talk about committing suicide from those who act on the impulse.

3a. On the basis of their expected high tendency to action and activity, the suicide committers were expected to communicate more in an unstructured communication situation than nonsuicidal individuals. Therefore, it was predicted that their actual suicide notes would be longer than those of nonsuicidal persons. (This hypothesis was not supported by the data.)

3b. It was hypothesized that suicidal persons are highly motor-action oriented. Verbs of action, which involve actual physical motor behavior rather than reflecting thought or feeling, were predicted to be found more often in the messages of the C/S group than the N/S group. (This prediction was supported in the findings. The T/S group differed from the C/S group relative to this variable, and was like the N/S group.)

3c. It was expected that the method of suicide attempt would be more active (based on ratings of skilled suicide clinicians) for suicide-committers than that stated by nonsuicidal individuals. (Support was not found for this hypothesis.)

3d. A second hypothesis related to activity which was supported in this study was that the individual notes of suicidal persons will, more frequently than the notes of nonsuicidal persons, contain a greater number of verbs than adjectives. This notion is based on Boder's finding that the adjective-verb quotient is indicative of whether an individual is more concerned with contemplation or with action. (In this index of activity, the C/S group was significantly higher than the N/S group. The T/S group was very different from the C/S group and more like the N/S group.)

4. *Expression of Anger*

4a. Because of their life style of covert aggression, the notes of suicide-committers were expected to contain self-accusations and thoughts that others have been hurt by their past behavior more frequently than the notes of nonsuicidal persons. (This hypothesis was not supported by the data.)

4b. Attack in the form of blame was predicted in the notes of actual suicides. (Concern regarding the blame of others for his suicide was found to be communicated more frequently in the C/S notes than the N/S

notes. The T/S group was more like the C/S than the N/S group regarding this hypothesis.)

4c. The use of "suffering messages" to influence others was expected to be prominent in suicidal communications. Reference to the communicator's suffering was expected to be found more frequently in notes of suicide-committers than those of nonsuicidal persons. (The findings showed no statistical difference between the groups in the frequency of this type item. All groups included a high number of these items.)

4d. The next prediction was that the suicidal individual would attempt to make sure his veiled aggression hits his targets. (This hypothesis was supported.) Messages related to the notification of a specific person or addressed to a special person were found more often in the suicide notes of the C/S group than of the N/S group. (The T/S group was like the N/S group and differed from the C/S group regarding this variable.)

4e. Content expressing "direct anger" was expected to be found less frequently in the notes of suicidal individuals than in those of nonsuicidal individuals. (This prediction was not supported statistically because there were too few expressions of direct anger in any of the notes to permit a statistical comparison of difference.)

4f. In keeping with Adlerian and psychoanalytic theory generally, it was predicted that suicidal individuals to a much greater degree than nonsuicidal persons would write notes whose content reflected a general tenor of veiled hostility, aggression, and resentment. (The prediction was supported by the data. The T/S group was more like the C/S group than the NS group in this respect.)

DISCUSSION

The findings of this study indicate that the T/S group is both similar and dissimilar to the C/S group and to the N/S group. A question arises whether the T/S group is from a single population, relative to the variable of suicide, or if it is possibly from two or more populations, varying in degrees of suicidality. If suicidality is considered as the degree of likelihood that an individual or group of individuals will successfully act out the impulse to commit suicide, it seems reasonable to

assume that the more suicidal a person is, the more he will have a communication style similar to the C/S group. If it were possible to split the T/S group into high-potential suicide threateners and low-potential suicide threateners, would the high-potential threateners communicate more like the C/S group and the low-potential threateners like the N/S group?

To answer this question, the suicide threateners were dichotomized into 11 high-potential suicide threateners and 11 low-potential suicide threateners, using the Los Angeles Suicide Prevention Center's Lethality Assessment scale (Hattem, 1963; Scheitzach, 1964). These two groups were then compared in terms of the communication-style items. Unfortunately, the *N* was so small that none of the items was able to differentiate the two groups to a statistically significant degree. All trends, however, except three which tied frequencies, were in the expected direction. The fact that not a single trend was counter to that which was predicted, suggests that larger sample comparisons might yield statistically significant differences.

The findings that the TS group shows a heightened importance of parent figures, a need for emotional support, less concern than the C/S about loss, and less need than the C/S to have things done for him, may offer some explanation as to what helps a person who comes for psychotherapy while threatening suicide, to work against his suicidal impulses. While the therapist becomes an important parent figure to whom the patient looks for support, the patient is able at some point to accept the responsibility for therapy and to value doing something for himself. This dynamic, along with ego strength to adapt to losses, helps him acquire a more constructive communication and life style.

Another specific finding of interest was that the direct expression of anger was almost totally absent from all notes, in all three groups. This suggests that in this culture there is great difficulty in consciously perceiving a relationship between anger and suicidal impulses, not only by the suicide-committer himself, but also by suicide-threatening and nonsuicidal individuals.

While such a step involves much conjecture

at this point, it seems worthwhile to posit some relationships between the variables which were found to be significantly related to suicide in this study. Such a consideration may offer insight into possible dynamics involved in the development of a personality prone to suicide. The data suggests that individuals who commit suicide may have an early history of being treated in such a way that they develop feelings that they need to be pampered by others. The predisposing experience of persons who display this strong need for pampering may be related to poor handling of dependency needs by the parents, where either excessive or deficient "feeding" (or an inconsistent combination of both) was the pattern. As a result of this uncertainty about whether his survival needs will be met, the individual retains his infantile self-centeredness and continuing question concerning his survival. The feeling grows that others do not really want to care for him, so if care is to be gotten, others must be manipulated into giving it. Success at manipulation requires the development of keen awareness of others in a self-centered way. Any seeds of a concept of social interest die out quickly as they go unfed. The individual learns to sense and use the unconscious guilt of others. Since his safety and survival needs are not adequately met, his actions often lead to hurt and painful experiences. A surface feeling of his importance to others results from his noting that others give him attention when he hurts himself. Since attention is as close to caring as the developing suicidal person thinks he can get, these experiences become very important ones for him. Attention becomes a substitute for the caring experience. The dependency of the suicidal person differs from the dependency found in the constellation of dynamics of numerous other emotional disorders. It is an active dependency, rather than a passive dependency. A passive-dependent personality sits and waits while others make food and other things available to him. The suicidal personality is active, but the activity is directed at manipulating others to meet his needs, rather than accomplishing things to satisfy his own needs. This pattern results in his being unable to develop any feelings of real individuation or self-worth, partly be-

cause he never accomplishes anything for himself, and partly because he never develops an inner feeling of safety which is necessary to experience himself as an individual. Unable to be comfortably independent, he sees others as the only source for getting his needs met. He is painfully aware at the same time that others are the frustrators of his needs. Hating others for this set of events, he is nonetheless afraid to express his anger directly for fear of losing what he does have. Thus, a repetitious pattern of hurting himself to use the guilt of others to hurt them and to manipulate them becomes subjectively the most reasonable style of coping.

Certain implications for diagnosis and treatment of suicidal persons may be drawn from the conclusions. It may be possible to detect individuals with highly self-destructive life patterns from an analysis of other message forms more readily available in clinical settings. Responses to open-ended written applications for psychotherapy or tape recordings of initial phone contacts are two obvious forms which could be objectively analyzed. Information from anamneses and projective tests could also be used to gain insight into the patient's life style and communication style. Treatment should be directed at helping the patient become aware of his intrapunitive active-dependent coping style by making objective reference to his parallel style and content of communication during the process of psychotherapy.

REFERENCES

- ADLER, A. Suicide. *Journal of Individual Psychology*, 1958, 14, 57-61.
- BATESON, G., JACKSON, D. D., & WEAFLAND, J. Toward a theory of schizophrenia. *Behavioral Science*, 1956, 1, 251-264.
- DARBONNE, A. R. Suicide and age: A suicide note analysis. *Journal of Consulting and Clinical Psychology*, 1969, 33, 46-50.
- FARBEROW, N. L., & SHNEIDMAN, E. S. Suicide and age. In E. S. Shneidman & N. L. Farberow (Eds.), *Clues to suicide*. New York: McGraw-Hill, 1957.
- FARBEROW, N. L., & SHNEIDMAN, E. S. (Eds.) *The cry for help*. New York: McGraw-Hill, 1961.
- HATTEM, J. The precipitating role of discordant interpersonal relations in suicidal behavior. Unpublished doctoral dissertation, University of Houston, 1963.
- MAINLAND, D., & MURRAY, I. M. Tables for use in fourfold contingency tables. *Science*, 1952, 116, 591-594.
- MENNINGER, K. *Man against himself*. New York: Harcourt, Brace, 1938.
- MOOD, A. M. *Introduction to the theory of statistics*. New York: McGraw-Hill, 1950.
- OSGOOD, C., & WALKER, E. G. Motivation and language behavior: A content analysis of suicide notes. *Journal of Abnormal and Social Psychology*, 1959, 59, 58-67.
- RUESCH, J. *Therapeutic communication*. New York: Norton, 1961.
- SCHMITZACH, J. An evaluative comparison of the dyadic relationship wherein one member engages in self-destructive behavior. Unpublished doctoral dissertation, University of California at Los Angeles, 1964.
- SHNEIDMAN, E. S., & FARBEROW, N. L. (Eds.) *Clues to suicide*. New York: McGraw-Hill, 1957. (a)
- SHNEIDMAN, E. S., & FARBEROW, N. L. Comparison between genuine and simulated suicide notes by means of the Mowrer DRQ. *Journal of General Psychology*, 1957, 66, 251-256. (b)
- SHNEIDMAN, E. S., & FARBEROW, N. L. The logic of suicide. In E. S. Shneidman & N. L. Farberow (Eds.), *Clues to suicide*. New York: McGraw-Hill, 1957. (c)
- SHNEIDMAN, E. S., & FARBEROW, N. L. A socio-psychological investigation of suicide. In H. P. David & J. C. Brangelman (Eds.), *Perspectives in personality research*. New York: Springer, 1960.
- SULLIVAN, H. S. *The interpersonal theory of psychiatry*. New York: Norton, 1953.
- TUCKMAN, J., KLEINER, R. J., & LAVELL, M. Emotional contents of suicide notes. *American Journal of Psychiatry*, 1959, 116, 59-63.
- TUCKMAN, J., KLEINER, R. J., & LAVELL, M. Credibility of suicide notes. *American Journal of Psychiatry*, 1960, 116, 1104-1106.

(Received September 16, 1968)

PREDICTING THE SUCCESS OF PEACE CORPS COMMUNITY DEVELOPMENT WORKERS¹

CHARLES DICKEN²

San Diego State College

Predictor assessments obtained before and during training of 55 Peace Corps Volunteers were compared with ratings of effectiveness of performance overseas. The main criterion was a composite of judgments by field supervisors; there were at least two raters for each S. The most valid predictors were peer ratings, life history, certain training grades, and ratings by a selection board using all data. Some psychometric measures of personality and cognitive style had modest validity, but intellectual measures had almost none. Interviews and clinicians' impressionistic interpretations of predictor data showed low validity. Females were more predictable than males. Analysis of a subsample of "least effective" Ss failed to identify predictors particularly effective in discriminating between candidates "just acceptable" and "just not acceptable." The predictors were factor-analyzed.

Mischel (1965) compared criterion success ratings of 35 Peace Corps Volunteer teachers in Nigeria with psychological assessments obtained while the Volunteers trained to go overseas. Self-report personality measures correlated significantly with the criterion but other predictors did not. The predictors employed and their correlations with the criterion were as follows: California F Scale (-.45), Barron Ego-Strength scale (.34), Taylor Manifest Anxiety scale (-.34), peer ratings (.29), faculty ratings (.27), academic grades (.24), assessment board (.20), and interview (.13). Mischel concluded that his results supported the usefulness of simple self-report devices and that they pointed, as have previous studies, to the limitations of "global" assessment procedures. Limitations in Mischel's predictor and criterion measures (frankly avowed by him) will be discussed later.

¹ This study was supported by a grant from the San Diego State College Foundation and also by the College Computer Facility. The data of this study were obtained with the assistance of the Peace Corps. The study, however, is the responsibility of the author and should not be construed as an official Peace Corps document. The author is indebted to the many colleagues who participated in data collection. Special thanks are due Roy McDonald, who participated extensively in obtaining and evaluating the predictive assessments.

² Requests for reprints should be sent to Charles Dicken, Department of Psychology, San Diego State College, San Diego, California 92115.

Ezekiel (1968) obtained fictional autobiographical essays describing the personal future from 42 Volunteers training to serve as teachers in West Africa. The essays were scored for *Differentiation*, degree of complexity and detail (three subscores); *Demand*, long-term effortful response to challenge; and *Agency*, portrayal of self as agent of major decisions. The scores' correlations with a composite first-year criterion success rating ranged from .13 to .31 and with a second-year rating by the Peace Corps representative from .22 to .34. The sum autobiography scores correlated with these criteria .29 and .41. There is some question of contamination of the success ratings by the predictors since the investigator and the training-site selection officer collaborated with the representative and his deputy in determining the first-year ratings. The autobiography measures had negligible validity in predicting peers' ratings of teaching success for the total sample, but did predict for female and Protestant subgroups (validity of autobiography sum .41 and .56). Barron's Ego-Strength, Taylor's *MA* scale, and an authoritarianism measure were obtained in training, but no validities were reported.

The present study investigated the validity of a comprehensive assessment program in predicting the success of 55 Volunteers who served as community development workers in Peru.

METHOD

*Assessment Program and Predictor Variables*³

At the time the Ss of the study entered Peace Corps training, applicants for the Peace Corps were initially screened by college-educated technicians under supervision of a clinical psychologist. In addition to skill classification, the technician made a *pretraining assessment* of overall suitability for the project to which the candidate was assigned. This was based on an application blank, educational transcripts and employment records, letters of reference, and scores on the *General Ability Test* and *Modern Language Aptitude Test*. The pretraining assessment was summarized in a short sketch for use at the training site. The technician also rated the candidate on the scale: (1) Unsatisfactory, (2) Marginal, (3) Good, (4) Very good, (5) Excellent. Plus (+) and minus (-) were used to qualify the numerical ratings in many instances. Candidates invited to train were all rated 3 or higher. During the study period approximately half of all applicants to the Peace Corps were invited to train and about half of all invitees accepted.

At the training site two clinical psychologists were employed as field assessment officers (FAO) responsible for obtaining and interpreting assessment data. Another clinical psychologist, the field selection officer (FSO), made the final decision of whether the trainee was sent overseas ("selected") or not sent ("deselected"). The FSO was assisted in his decisions by a selection board composed of himself, the FAOs, the training project director, the project psychiatrist, a training officer from Peace Corps headquarters, and a representative of Peace Corps administration in Peru. The selection board met at the midpoint of the 12-week training program and again at the close.

A description of the quantitative or quantifiable assessment data gathered during training follows. The listing approximates the sequence in which the data were obtained, with exceptions noted.

Test data. The Minnesota Multiphasic Personality Inventory (MMPI) and the *Crutchfield Figures* were administered at the start of the training. The MMPI was scored on the standard 13 scales plus the Barron Ego-Strength (*Es*) scale.

The senior FAO, who was experienced in the clinical use of the MMPI, made a "blind" *MMPI rating* and wrote an interpretation for the selection board. Prior research has indicated low validity for the MMPI clinical scores as predictors of work performance of normals (Dickens & Black, 1965). Only *Es* was analyzed separately. The Crutchfield is an embedded figures series aimed at measuring perceptual and cognitive flexibility.⁴ The FAOs also had access to the application tests cited above.

Initial FAO impressions. During the first weeks of training, the FAOs conducted a structured interview with each trainee. Emphasis was on assessing motivation, maturity and emotional stability, interpersonal relationships, leadership, competence, and psychiatric

difficulties, if any. The *interview rating* of overall potential was then done with knowledge of the pretraining assessment and test data.

Near the midpoint of training, a leaderless *discussion group* session was held with sets of five to seven same-sex trainees. Each group was asked to attempt to solve the problem of an "unsuitable Volunteer" overseas. The FAOs observed and rated overall potential as indicated by discussion behavior.

Throughout the training, the FAOs made observations of performance and social behavior.

Training grades. The study uses grades made at the end of the program by the staff of four training components. *Community development* training performance was graded by examinations, discussion, and field work observations. The *language rating* was based on the Foreign Service Institute structured oral examination in Spanish. *Physical education* grades reflect attitude and effort more than athletic proficiency per se. *Domestic skills training* was primarily for the women trainees, but the males were given brief training and both sexes were graded. Male technical skills were not graded. The training faculty made qualitative appraisals of personality which were available to the selection board.

Peer nominations. At the close of the program, each trainee was asked to nominate five others of each sex he would choose to work with overseas (*assignment preference*), five of each sex with the most *leadership* potential, and any number he believed had *overall potential* for success in Peace Corps work. Two other nominations "adapt well overseas" and "can confide in" were obtained but not analyzed for the present study because of high intercorrelations with other ratings. An unlimited number of *negative nominations* was permitted on each of the five items (nominator prefers not to be assigned with S, sees lack of leadership potential in S, etc.). There were 65 nominators. Descriptive comments by peers were obtained and used by the selection board.

Final FAO impressions. These were based on all data available to the FAOs during the training process but were not recorded until after the action of the final selection board. *Board status* was the FAOs' impression of the intermediate and final selection board discussions of the trainee including controversy. *Development* was based on presence in the life history of "crisis experiences" and successful resolution of same. *Overall potential* was the FAOs' global impression of the trainee.

Final selection board rating. Trainees selected for overseas service were rated 4 ("is approximately in the lower quartile"), 5 ("approximately the middle two quartiles"), or 6 ("approximately the top quartile"). The psychiatrist contributed observations to the board based on all trainees' performance in mental hygiene training and on interviews conducted with about one-fourth of the study sample who had been referred for evaluation. All assessment data was available to each board member. The FSO also used the results of a Civil Service background investigation of each trainee.

Gough MMPI scales. Thirteen of the special MMPI scales developed by Gough (Dahlstrom & Welsh,

³ Variables used in the present study are italicized.

⁴ R. Crutchfield, personal communication, June, 1964.

1960, pp. 448-467) were scored for the present study. These scores were not available during the assessment process and for this reason were analyzed separately from the other predictors.

Subjects

Eighty-four candidates accepted invitations to join the 12-week training project. Resignations and deselection for suitability or medical reasons left 61 candidates who were sworn in as Volunteers at the close of training and sent to Peru. Criterion data became available for 55 of the 61, which constitute the total sample of the study. The median age during the training period was 22. Two Ss had completed bachelor's degrees and some postgraduate education, 48 had completed the bachelor's, four 2 years of college, and one a single year of college.

Performance Criteria

The Volunteers were, in the main, assigned to work in *barriadas*, relatively new owner-constructed, quasi-legal settlements which have grown up outside established cities in the coastal regions of Peru. The task of the Volunteer was to stimulate and organize community action projects in housing, roads, education, youth activities, health, and related areas. An important component of the task was to foster communication between the *barriada* dweller and agencies of the Peruvian government. The job of the community development worker has been described as "unstructured" in that there is usually no clearly specified task for the worker when he arrives in the community. Results are judged as much in terms of increased cohesiveness in the community as by specific projects completed. Providing technical assistance and instruction was frequently part of the task but usually not the main emphasis.

Composite field rating. The main criterion was an average of independent ratings of Peace Corps administrators (associate representatives and deputy director) responsible for supervising the Volunteers' work. The raters judged Ss on "Overall effectiveness as a Volunteer, taking site difficulty into account." A 7-step scale was used: (1) Poor, (2) Below average, (3) Probably below average, (4) Average, (5) Probably above average, (6) Above average, and (7) Superior. The raters were instructed to use their total experience with Volunteers in setting a frame of reference for "Average," "Poor," etc. A standard of at least two raters per S was set. Fifty-one Ss (23 male, 28 female) qualified for this criterion measure. The ratings were made after 12 to 16 months' service by S. Nine raters participated and contributed a total of 134 ratings. There was an average of 2.9 raters per male S and 2.4 raters per female S. Reliability estimates for the composite ratings determined by Ebel's (1951) technique are .71 (males), .50 (females), and .62 (total). The reliability estimates were computed for the case in which the source of the ratings was not the same for all Ss, and were functions of the number of raters per S and the degree of interrater agreement. The amount of contact between rater and S varied con-

siderably. No S was rated, however, unless the rater felt he had an adequate basis for determining the criterion as specified.

The composite field rating on the 51 Ss is referred to hereafter as the "main" criterion. Two alternate criteria were available for some of the 51 Ss and in one instance for two additional Volunteers from the training project.

Researchers' rating. Two staff members of the Peace Corps research division briefly visited 30 Volunteers and their supervisors. Two of the Volunteers were not among the 51 on whom composite field ratings were obtained. Fifteen of the 30 were males, 15 females. The Volunteers had served 10 months at the time of the visit. The researchers made a single performance rating of each S and later made these available to the author.

Inspector's rating. The second alternate criterion was furnished by an inspector charged with evaluating training and field programs, who visited 21 of the Ss after approximately 10 months' service. Subsequent to furnishing these ratings, the inspector became deputy director in Peru and contributed to the composite field ratings.

The alternate criteria were based on fewer Ss and briefer contact than the main criterion. Their reliability cannot be estimated directly. The researchers' rating correlates .42 with the composite criterion ($N=28$) and the inspector's rating correlates .66 ($N=20$). The alternate criteria intercorrelate .64 ($N=19$).

In addition to the 53 Ss on whom performance ratings were available, two additional Ss from the original 61 were analyzed solely with regard to early termination (described later).

Analysis

Means, ranges, standard deviations, and Pearson intercorrelations of the variables were computed. N is maximized in each analysis; thus intercorrelations among the predictor variables uses $N=53$; correlations between these variables and the composite field rating criterion uses $N=51$, etc. A tri-serial correlation was computed between the final selection board rating and the main criterion. A factor analysis was done on the 20 predictor variables. The solution was reached by principal axes with unity in the diagonals and a Varimax rotation. Kaiser's (quoted in Cooley, 1962, p. 160) criterion for the number of factors (no latent roots less than unity) was applied and resulted in extraction of six factors.

Canonical (multiple) correlations between selected predictors and the main criterion were computed. Selection of predictors was done with several strategies: (a) all predictor variables; (b) best representation of each class of data (pretraining data, test scores, initial FAO impressions, grades, peer ratings, final FAO impressions); (c) factor representation (most valid representative of each factor); (d) economy (use of measures available without psychologists' effort), and (e) simplicity (pretraining assessment and peer leadership rating).

Possible Influence of the Criteria by the Predictors

At the time of the study, detailed assessment data from the training site were not furnished to administrators overseas. The selection board rating and a general, favorably-toned résumé of the Volunteers' skills and characteristics were forwarded, and the overseas representative to the selection board had an opportunity to confer with some of the administrators at the time the Volunteers were assigned. Communication difficulties within the host country and inefficiencies in forwarding data sometimes meant no training data reached the administrators.

The sparseness of available training data, the length of time between the arrival of the Volunteers in Peru and the point when the criteria were obtained, and the well-known pragmatism of Peace Corps administrators suggest that the field ratings were based on the Volunteers' actual performance and influenced little if at all by information produced at the training site. A notable possible exception is the field data furnished by the representative (G.B.) who served on the selection board and who later contributed 19 (14%) of the 134 main criterion ratings. Two analyses were done to evaluate the plausibility of the assumption that the predictor data had little influence on the criteria and to investigate the exception noted.

Survey of the raters. Seven of the eight main criterion raters other than G.B. were queried about their access to training data and their impression of the influence of such data on their field ratings.⁵ Three of the seven, who furnished 18 (13%) of the ratings, stated they had no access to training data. A rater who made 12 (9%) of the ratings stated, "the feedback from training was too slow. We had very little data until the Volunteers were assigned. Unfortunately, we had to rely too heavily on the biographical sketch."⁶ A representative who made 4 (3%) of the ratings said, "I do not recall what [training data] I had . . . prior to my last assessment, but I doubt that it influenced that assessment." A representative who furnished 14 ratings (10%) stated,

If there was [influence] it wouldn't be very much. Usually 16 months is far enough—you pretty much go on what the Volunteer has done. I would say bias was minimal. You learn first impressions aren't worth much. "Gold Stars" would come in, and they would not turn out—and vice versa.

The deputy director, who furnished 45 ratings (34%), and who had earlier done the inspector ratings, responded as follows,

I was influenced by nothing, to the best of my knowledge, other than the Volunteers' performance

⁵ This information was obtained after the ratings were received. One rater proved impossible to reach; he had contributed 22 (16%) of the main criterion ratings.

⁶ A short paragraph, neutral or favorable in tone, composed entirely from information furnished by the Volunteer.

in Peru, in filling out the ratings you requested. The only data . . . that figured in my judgment was my own observation and [the] immediate field supervisors' evaluations. I presumably had seen some training data on them, but I can assure you it had no influence on my ratings.

Comparative predictability of criterion ratings made by "most influenced" rater. Since rater G.B. had served on the selection board before assuming his administrative post in Peru, his field ratings would presumably be most open to contamination by the training data. Not only did he have complete access to all the predictors, numerical and qualitative, but he was also personally involved in the assessment process. Criterion contamination associated with this rater would, if present, be expected to result in greater predictability of his field ratings than of the overall composite field ratings, since the latter were based mainly on judgments by raters exposed to considerably less training data and who had no personal involvement in the assessment and selection of the Volunteers. To test this, G.B.'s field ratings were separately correlated with the final selection board rating and with the three best individual predictors. The outcome of this comparison appears in the Results and Discussion section.

Since many of the raters did have access to at least some information from the training site, despite the time lapse, the raters' denial of influence, and the evidence from the comparison cited above, the possibility of some criterion contamination cannot be entirely ruled out. The findings of the study are subject to this limitation. Since no raters except G.B. had access to the detailed numerical assessments that are used as predictors, it seems probable that contamination, if present, has relatively little influence on comparison of the relative efficiency of the various predictors, even though overall predictive validity might be inflated to some degree.

RESULTS AND DISCUSSION

Table 1 shows scale properties⁷ of the major variables and the ranges, means, and standard deviations for all Ss and for males and females separately. Range is computed by (upper limit–lower limit); thus the maximum possible range for a 7-step scale is six units. *T* scores were based on Peace Corps applicant norms except for *Es*, where Barron's norms were used.

The variabilities are in most instances substantial. The ranges are often at or near the maximum possible. Thus restriction of range is in most instances not a factor restricting correlation between variables. Males vary more on the criteria than females and tend to do so on the predictors. There is no con-

⁷ Some of the original scales (e.g., the pretraining assessment ratings) were expanded to accommodate "qualified" ratings ("3+," "3/4," etc.).

TABLE 1
PROPERTIES OF VARIABLES

Variables	Scale	Steps in scale	Total (N = 53)			Males (N = 24)			Females (N = 29)		
			Range	M	SD	Range	M	SD	Range	M	SD
Criteria*											
Composite field rating	rating	7	5.4	4.1	1.4	5.4	4.2	1.5	4.3	4.1	1.3
Researchers' rating	rating	5	4	2.8	1.1	4	3.2	1.0	3	2.3	1.0
Inspector's rating	rating	9	8	4.9	2.4	8	5.4	2.3	6	4.2	2.3
Predictors											
Pretraining assessment	rating	6	4	3.9	.9	4	3.9	1.0	4	3.9	.9
General Ability Test	T score	—	31	57.7	7.1	30	59.6	6.4	30	56.1	7.2
Modern Language											
Aptitude Test	T score	—	29	57.1	7.4	27	55.4	8.0	24	58.5	6.4
Crutchfield Figures	raw score	21	20	11.2	5.2	20	13.2	5.4	18	9.6	4.5
MMPI evaluation	rating	5	4	3.6	1.3	4	3.4	1.3	4	3.8	1.3
MMPI Es	T score	—	41	65.7	7.3	31	64.7	6.9	39	64.0	7.6
Interview	rating	8	7	5.2	1.8	5	5.6	1.6	7	5.0	1.9
Group discussion rating	rating	7	6	4.8	1.7	6	4.9	1.6	6	4.6	1.8
Community development grade	rating	8	7	6.0	1.8	7	5.8	2.1	5	6.1	1.5
Language rating	rating	8	7	3.3	1.7	7	3.4	1.5	7	3.2	1.8
Physical education grade	rating	6	5	4.1	1.1	5	4.3	1.1	4	4.0	1.1
Domestic skills grade	rating	5	2	4.5	.6	2	4.5	.6	2	4.4	.6
Peer assignment preference	raw score	64	23	9.4	4.8	20	11.0	4.9	18	8.1	4.4
Peer leadership nomination	raw score	64	47	10.1	11.9	47	12.6	12.9	42	8.0	10.4
Peer overall nomination	raw score	64	42	10.3	10.1	42	12.8	11.2	38	8.2	8.5
Peer negative nomination	raw score	321	29	4.2	5.6	14	3.6	3.9	29	4.5	6.6
FAO rating "board status"	rating	6	5	4.5	1.5	5	4.7	1.6	4	4.4	1.3
FAO rating "development"	rating	8	6	5.4	1.7	6	5.4	1.5	6	5.3	1.8
FAO rating "overall potential"	rating	8	6	5.6	1.8	5	5.9	1.8	6	5.4	1.8
Final selection board	rating	3	2	4.8	.7	2	4.8	.7	2	4.7	.8

*Ns for criteria (total, Male, Female): Composite (51, 23, 28), Researcher (30, 15, 15), Inspector (21, 13, 8).

sistent sex trend in mean values except in peer ratings where the males excel.

Table 2 shows the intercorrelations of the predictors. The .05 and .01 significance levels for correlations are .27 and .35 ($df = 51$). A factor analysis of the predictors is presented later. Correlations and factor loadings are shown in the tables with decimals omitted.

Predictive Validities

Table 3 shows predictive validities of the assessments (r_{pc}) and validities corrected for unreliability of the criterion (r_{pcor}). The coefficients are shown for males and females separately in the case of the main criterion. The most effective predictors of the main criterion were the pretraining assessment, community development grade, and peer leadership rating. Each equals or exceeds the

validity of the final selection board which used all data. Test scores and psychologists' impressions have relatively low validity.

The uncorrected correlations of the selection board rating and the three best individual predictors with the criterion ratings by representative G.B. ($N = 19$) were *final selection board*, .09; *pretraining assessment*, .32; *community development grade*, .35; and *peer leadership nomination*, .36. Comparison of these correlations with those in the first column of Table 3 indicates that in each instance the outcome is opposite to that predicted by the contamination hypothesis. Failure of the final selection board ratings to predict G.B.'s field ratings indicates his impressions of the Volunteers' field performance deviated quite distinctly from the appraisal to which he had been both privy and a contributor 16 months

TABLE 2
INTERCORRELATIONS AMONG PREDICTORS

Predictor	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Pretraining assessment	21																			
2. General Ability Test		37																		
3. Modern Language Aptitude Test		25	57																	
4. Crutchfield Figures		25	53	34																
5. MMPI evaluation		02	-11	13	11															
6. MMPI E _s		27	26	38	01	22														
7. Interview		40	26	16	-04	04	41													
8. Group discussion rating		19	02	12	-10	05	27	41												
9. Community development grade		18	04	34	-13	-05	18	25	34											
10. Language rating		15	05	12	-12	01	08	34	-01	31										
11. Physical education grade		19	00	10	22	03	07	16	31	28	29									
12. Domestic skills grade		19	03	10	-14	00	10	21	16	13	27	33								
13. Peer assignment preference		19	12	04	06	07	15	33	21	33	46	21	24							
14. Peer leadership nomination		41	14	33	16	12	26	38	29	46	55	41	29	53						
15. Peer overall nomination		37	10	21	08	11	18	40	29	45	62	33	33	70	87					
16. Peer negative nomination		-04	09	09	07	-09	-01	14	01	-04	-08	08	03	-33	-12	-21				
17. FAO rating "board status"		16	24	26	06	-04	34	52	49	47	38	34	41	43	58	57	-04			
18. FAO rating "development"		28	22	20	03	04	31	49	40	26	42	19	36	33	53	-22	76			
19. FAO rating "overall potential"		22	30	26	06	-03	24	53	43	37	54	29	31	46	63	64	82	87		
20. Final selection board		29	20	23	03	13	22	56	44	50	53	30	39	52	65	72	78	79	83	

Note.—Total sample $N = 53$.

earlier. The data appear to justify the inclusion of G.B.'s ratings in the criterion and lend weight to the testimony of the raters and other considerations suggesting little criterion contamination.

Except for the MMPI evaluation, females tend to be more predictable than males (conventional wisdom notwithstanding!). This is true both of the level of the coefficients and in the larger number of assessments which show some validity. When the lower criterion reliability of the female sample is considered, the effect is heightened. There is other evidence of greater female predictability (Seashore, 1962).

Table 4 shows the multiple correlations obtained by various strategies of combining predictors. Uncorrected correlations were used in calculating the multiple correlation (R). Combinations of easily obtained data predict better than the selection board and much better than the assessment officers' impressionistic combination of data. Multiple correlations capitalize on sampling fluctuations in the correlations, and there is no way to estimate the shrinkage which would occur on cross-validation. This limitation is especially pertinent to the 16-predictor coefficient.

Clinical and Actuarial Prediction

From the pragmatic viewpoint, the most significant finding is that the pretraining assessment is virtually as valid as anything else, and that a combination of it with a single, easily obtained, in-training assessment (peer leadership) bests the validity of the selection board. Peer ratings and the pretraining assessment, while "clinical" in character, are routine and relatively inexpensive. The generally poor showing of impressionistic modes of combining data is certainly nothing new (Meehl, 1954; Sawyer, 1966).

The alternate criteria are somewhat better predicted by the psychologists' impressions and the selection board than is the main criterion. While the small N 's preclude any but tentative generalizations, there is the possible implication that a certain "superficiality" of the quickly obtained alternate criteria is shared by the psychologists' predictive impressions.

TABLE 3
CORRELATIONS BETWEEN PREDICTORS AND PERFORMANCE CRITERIA

Predictors	Composite field rating						Research- ers' rating	Inspec- tor's rating
	Total (N = 51)		Males (N = 23)		Females (N = 28)		15M, 15F (N = 30)	13M, 8F (N = 21)
	r_{po}	r_{pco}	r_{po}	r_{pco}	r_{po}	r_{pco}	r_{po}	r_{po}
Pretraining assessment	41**	52**	37	44*	44*	62**	21	38
Test data								
General Ability Test	17	22	-07	-08	37	52	32	-02
Modern Language Aptitude Test	11	14	04	05	23	32	-13	-23
Crutchfield Figures	31*	39**	19	23	44*	62**	05	09
MMPI evaluation	14	18	44*	52*	-11	-16	00	27
MMPI Es	32*	41**	24	29	48**	68**	17	24
Initial FAO impressions								
Interview	23	29*	09	11	34*	48**	30	28
Group discussion rating	23	29*	22	26	24	34	20	38
Training grades								
Community development	43**	55**	41*	49*	49**	69**	11	27
Language rating	17	22	-03	-04	32	45	18	16
Physical education	32*	41**	12	14	50**	71**	07	24
Domestic skills	12	15	-13	-15	33	47*	48**	45**
Peer nominations								
Assignment	22	28*	22	26	21	30	40*	30
Leadership	50**	64**	48*	57**	51**	72**	35	56**
Overall	40**	51**	31	37	52**	72**	44*	60**
Negative	-20	-25	-13	-15	-25	-35	02	05
Final FAO impressions								
"Board status"	27*	34*	06	07	47*	66**	46**	43*
"Development"	26	33*	19	23	33	47*	51**	50*
"Overall potential"	22	28*	13	15	30	42*	48**	51*
Final selection board rating (Triserial)	37**	47**	24	29	47*	66**	53**	51*
	41**	52**	27	32	55**	78**	—	—

* $P_r < .05$.** $P_r < .01$.

Prediction of Least Successful Performance

The function of the assessment program and of the psychologists' and selection board's judgments was not primarily to predict degree of success among suitable prospects but to separate the unsuitable. The "band" to which the assessments were oriented was "barely not suitable" to "barely suitable." Since individuals found unsuitable were not measured as to overseas performance, there is no direct measure of the accuracy of this discrimination.

An indirect test of the effectiveness of the assessments in this band is shown in Table 5. Means and standard deviations of selected predictors and significance tests are shown for 12 "least successful" Ss versus the remainder ("more successful"). The least successful con-

sist of those five males and five female Ss receiving composite (main) criterion ratings in the range 1.0-2.99 ("poor" to the upper limit of "below average") plus two additional early termination Ss. Six of the original 61 Volunteers terminated service prematurely. The official "reason" for termination and anecdotal data from field raters were available for all six. Three of the six terminators met the standard for the main criterion and are in the study sample. All of these received composite criterion ratings less than 2.99, and so fall in the "least successful" group. While confidentiality forbids going into detail, it seemed clear to the writer that two additional Volunteers (one male, one female) terminated because of, or concomitant with, unsatisfac-

TABLE 4
MULTIPLE CORRELATION OF PREDICTORS AND MAIN CRITERION

Strategy	Predictors	Number of predictors	R
All predictors	All shown in Table 2 except final FAO impressions and final selection board	16	.82
Best representation of classes of data	Pretraining assessment; <i>Es</i> ; interview; community development grade; peer leadership rating	5	.62
Factor representation	Peer leadership rating; Crutchfield Figures; FAO board status; PE grade; <i>Es</i> , pretraining assessment	6	.61
Economy	Pretraining assessment; community development grade; language grade; PE grade; peer leadership rating	5	.61
Simplicity	Pretraining assessment; peer leadership rating	2	.54

Note.— $N = 51$.

tory performance. These two are included in the "least successful" group. Five of the six early terminators are thus included in the least successful group. The other early terminator appears to have performed satisfactorily until leaving.

Were the assessments especially effective in the band at which they were aimed, it would be reasonable to expect relatively large differences in the means of the two groups relative to variabilities. This is by no means uni-

formly the case, and only 5 of 18 comparisons reach more than borderline significance.⁸ The

⁸ Because of the small N , the possibility of Type II errors must be considered, especially in the sex-separated analyses. On the other hand, the variance homogeneity assumption for t is probably not justified in the case of the peer and selection board data. Despite the large differences, especially in the peer nominations, nonparametric analyses for the total sample (Mann-Whitney) and for the females (Kruskal-Wallis) were not significant for either variable.

TABLE 5
LEAST SUCCESSFUL VOLUNTEERS VERSUS MORE SUCCESSFUL VOLUNTEERS: SELECTED PREDICTORS

Predictors	Total			Males			Females		
	Least ($N = 12$)	More ($N = 43$)	P_t	Least ($N = 6$)	More ($N = 19$)	P_t	Least ($N = 6$)	More ($N = 24$)	P_t
Pretraining assessment									
<i>M</i>	3.58	4.00	<i>ns</i>	3.67	4.00	<i>ns</i>	3.50	4.00	<i>ns</i>
<i>SD</i>	.76	.94		.75	1.03		.76	.87	
MMPI evaluation									
<i>M</i>	3.42	3.72	<i>ns</i>	2.33	3.79	.05	4.50	3.67	.10*
<i>SD</i>	1.44	1.28		1.11	1.10		.76	1.40	
Interview									
<i>M</i>	5.00	5.23	<i>ns</i>	5.33	5.58	<i>ns</i>	4.67	4.96	<i>ns</i>
<i>SD</i>	1.83	1.78		1.37	1.60		2.13	1.86	
Community development grade									
<i>M</i>	5.17	6.19	.10	5.33	6.95	.10	5.00	6.38	.10
<i>SD</i>	1.62	1.79		1.70	2.16		1.53	1.41	
Peer leadership nomination									
<i>M</i>	5.42	11.09	.05	8.17	13.68	<i>ns</i>	2.67	9.04	.05
<i>SD</i>	5.69	12.63		5.49	14.01		4.42	11.00	
Final selection board									
<i>M</i>	4.42	4.81	.05	4.67	4.84	<i>ns</i>	4.17	4.79	.05
<i>SD</i>	.49	.79		.47	.74		.37	.82	

* Direction contrary to hypothesis.

TABLE 6
FACTOR MATRIX FOR PREDICTORS

Variable	Factor						
	I ^a	II	III ^a	IV	V	VI	k ^a
Pretraining assessment	28	31	05	03	11	52	46
General Ability Test	04	82	26	-10	-19	16	82
Modern Language Aptitude Test	-07	59	09	08	-01	57	69
Crutchfield Figures	06	84	-10	14	17	-06	78
MMPI evaluation	06	00	-02	-01	89	09	81
MMPI Es	-09	12	37	-08	33	59	57
Interview	20	-02	61	-12	06	38	57
Group discussion rating	-03	-24	55	24	06	44	61
Community development grade	44	-17	13	15	-30	61	72
Language rating	71	-02	23	11	-11	-03	59
Physical education grade	33	05	11	74	04	08	67
Domestic skills grade	19	-07	43	49	10	-13	50
Peer assignment preference	74	00	23	-09	07	06	62
Peer leadership nomination	74	15	30	22	09	28	80
Peer overall nomination	84	06	31	12	08	18	87
Peer negative nomination	-39	13	-12	61	-20	12	61
FAO rating "board status"	32	07	79	21	-11	18	83
FAO rating "development"	31	12	85	-01	03	03	83
FAO rating "overall potential"	44	15	80	07	-11	06	88
Final selection board	56	04	71	04	05	14	85

Note.—Total sample $N = 53$.

^a Signs reflected.

findings of considerable overlap in mean predictor scores do not change when the six cases with main criterion scores of 2.00 or less are compared with the remainder of the Ss or when the five terminated-unsatisfactory cases are compared with the remainder. Inspection does, however, indicate that the overlap is partly a result of two or three cases in the least successful group scoring quite high on predictors while the others in the group score low.

Inspection of the negative aspects of the assessment data on the least successful Ss yields the following impressions. Males: One of the six was assessed as essentially sound across the board and given a "5" rating by the selection board. A "clean miss" must be entered here. Of the other five, three had MMPI profiles rated "unfavorable" (Codes 4/9; 132'; and 83' with sharply elevated L and K) and one's profile was rated doubtful (59'). Two of the males showed prominent signs of anxiety in the interview plus psychosomatic illness in the life history. Two others were seen as strong, active individuals but with evidence of opinionation and excessive

individuality. All four last-cited males were seen by the selection board as having compensating strengths. The last of the five was perceived as immature and possibly lacking in staying power in pretraining assessment, community development training, and in the program generally. He was "controversial" at the midtraining selection board.

Females: One of the six was assessed as essentially sound. Three of the other five were seen during training as rather clearly lacking in experience and maturity, as naive, and as having questionable staying power. Likeability and inoffensiveness seems to have resulted in favorable selection decisions in these cases. The other two females were seen as stable and mature but as somewhat slow and passive during training.

In retrospect, there seems to have been a relatively clear basis for deselecting the one "poorest" male and the three "immature" females. The remaining eight least successful cases do not seem to have been clearly distinguishable from many other "selected" Ss because of what seemed, at the time, to be substantial compensating assets.

TABLE 7
CORRELATIONS BETWEEN GOUGH SCALES
AND MAIN CRITERION

Scale	No. of Items	Total (N = 50)	Males (N = 23)	Females (N = 27)
Academic achievement (Ac)	18	-.03	.00	-.05
Delinquency (De)	12	-.19	-.32	-.08
Dominance (Do)	28	.27*	.13	.41*
Femininity (Fe)	16	-.06	-.09	.07
Honor point ratio (Hr)	16	.23	.11	.34
Intellectual efficiency (Ie)	39	.04	.12	-.01
Impulsivity (Im)	21	-.15	-.04	-.30
Psychological interests (Py)	6	.25	.11	.44*
Social responsibility (Re)	32	.19	-.02	.45*
Social participation (Sp)	25	.21	-.02	.46*
Social presence (Sr)	21	.30*	.11	.44*
Social status (St)	34	.12	-.05	.22
Tolerance (To)	30	.31*	.12	.45*

* $P_r < .05$.

Factor Analysis of Predictors

The results of factor analysis of the predictors are shown in Table 6. The first factor obtained reflects overall merit, and loads most on peer ratings. Factor II is "tested mental ability." Factor III is equally clearly an "assessment officer" factor. With validity data in mind, the high loading of the selection board on this factor may document an undue influence of the FAOs on the board. Interpretation of the fourth factor is uncertain. Factor V is an MMPI factor. The last factor seems to reflect personal strength as assessed by life history, social workers (community development), psychologists, and psychological tests, but not by peers.

Validity of the Gough Scales

Table 7 shows the predictive validities of the Gough scales. Only Dominance, Social Presence, and Tolerance are significant in the total sample. None of the scales, with the possible exception of Delinquency, shows promise in assessing males. Several measures of social attitudes and behaviors show substantial validity for the females.⁹

Comparison with Mischel's (1965) Findings

The self-report personality measure (E_s) common to both studies yields similar validity

⁹ One female S was inadvertently omitted from the analysis. It should be pointed out that the Gough scales used here are not the longer versions containing specially prepared items which Gough later developed to measure the same traits in his California Psychological Inventory.

in each even though the measure in Mischel's study had the possibility advantage of being taken under "nonselection" (research) conditions. The present findings support Mischel's optimism over self-report measures only in the case of females. (Mischel's S_s were two-thirds male.) The interview yields similarly poor validity in both studies. Academic grades (at least the "better" ones), peer ratings (ditto), and selection board ratings yielded considerably higher validities in the present study.

Higher validity in the present study may result from several factors: (a) possible contamination of the criterion in the present study; (b) better criterion measures in the present study—Mischel's criterion did not permit a reliability estimate and was admittedly crude; (c) Mischel's predictive data used crude faculty and peer ratings, in that a composite was used in each case that may have pooled the effective with the ineffective; (d) "hurried" conditions under which Mischel's predictors were obtained; and (e) preponderance of males in his sample.

Factors which might have favored higher validity in the Mischel study are (a) the considerably more intangible nature of the criterion performance in the present study and (b) less restriction of range of talent by deselection and other attrition during training in the Mischel study (only 4 of 41).

REFERENCES

- COOLEY, W., & LOHNES, P. *Multivariate procedures for the behavioral sciences*. New York: Wiley, 1962.
- DAHLSTROM, W., & WELSH, G. *An MMPI handbook*. Minneapolis: University of Minnesota Press, 1960.
- DICKEN, C., & BLACK, J. Predictive validity of psychometric evaluations of supervisors. *Journal of Applied Psychology*, 1965, 49, 34-47.
- EBEL, R. Estimation of the reliability of ratings. *Psychometrika*, 1951, 6, 407-424.
- EZEKIEL, R. The personal future and Peace Corps competence. *Journal of Personality and Social Psychology*, 1968, 8(2, Pt. 2).
- MEEHL, R. *Clinical versus statistical prediction*. Minneapolis: University of Minnesota Press, 1954.
- MISCHEL, W. Predicting the success of Peace Corps volunteers in Nigeria. *Journal of Personality and Social Psychology*, 1965, 1, 510-517.
- SAWYER, J. Measurement and prediction, clinical and statistical. *Psychological Bulletin*, 1966, 66, 178-200.
- SEASHORE, H. Women are more predictable than men. *Journal of Counseling Psychology*, 1962, 9, 261-270.

(Received September 16, 1968)

COMPARISON OF THE BINET AND WPPSI WITH LOWER-CLASS FIVE-YEAR-OLDS

JOEN FAGAN,¹ ELIZABETH BROUGHTON, MILDRED ALLEN, BETTY CLARK

AND PATRICIA EMERSON

Georgia State College

The Stanford-Binet and WPPSI were administered to thirty-two 5-year-old lower-class children, 16 of whom were Negro and 16 white. A correlation coefficient of .80 was found between the Binet and WPPSI. Contrary to previous research, significant differences were found in the IQ scores obtained from the measures, with the Binet mean of 95.2 being eight points higher than the WPPSI. No significant sex or race differences were found in IQs. A decided preference for the Binet was reported by the examiners on the basis of ease of administration, greater scoring clarity, and subjective assessment of accuracy.

Increased emphasis on early education, especially the attempts to reach culturally disadvantaged and Negro children 1 or more years prior to the beginning of public school, has increased the need for adequate intellectual assessment instruments. The Stanford-Binet has been criticized for lack of Negro Ss in its standardization sample and its "middle-class" orientation which may penalize lower-class Ss. An extensive study aimed at providing norms for Stanford-Binet (Form L-M) performance by Negro school children in the Southeast states yielded a mean IQ of 80.6 (Kennedy, Van deRiet, & White, 1961). The publication of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) in 1966 made available another instrument which ostensibly corrects for the above faults. As yet there are few published studies which compare the WPPSI with other individually administered intelligence tests. The only published reliability and validity studies are found in the WPPSI manual (Wechsler, 1963). The study reported in the manual which correlated the WPPSI with the Stanford-Binet was administered to 98 children in a single elementary school in California. The mean Full Scale IQ obtained from the WPPSI was 89.6, and the Binet mean IQ was 91.3 (Wechsler, 1963, p. 33).

The present study was undertaken to com-

pare the WPPSI with the Binet with a group of lower-class Negro and white children. In addition to a statistical analysis, the Binet and WPPSI were compared on a subjective level for ease of administration, scoring, and S interest.

METHOD

The Ss were 32 children, between the ages of 4 years, 11 months, and 5 years, 11 months, attending day-care centers operated for children of working mothers living in low-income housing projects. Of these, 16 were Negro and 16 were white, with 8 boys and 8 girls in each of the two racial subgroups. The age range selected was such that the children would enter the first grade the following year. The day-care centers included two that were primarily white and four primarily Negro.

The tests were administered by four graduate students in psychology who had had prior experience with tests used in this study. Each child was given the two tests in two sessions from 1 to 5 days apart. The order of administration of the WPPSI and the Binet was counterbalanced. Standard procedure was observed in the administration of the tests with one exception. The Geometric Design subtest was found to be so time-consuming as to extend the administration of the WPPSI to two sessions, and the scoring was very difficult and subjective. Hence, the subscale was omitted and resulting performance IQs were prorated.

RESULTS

The means, standard deviations, and differences between means for the WPPSI and Binet are shown in Table 1. The highest mean IQ (95.2) was on the Binet. The Verbal, Performance, and Full Scale IQs of the WPPSI were almost identical, being between 87.1 and

¹Requests for reprints should be sent to Joen Fagan, Department of Psychology, Georgia State College, 33 Gilmer Street, S.E., Atlanta, Georgia 30303.

TABLE 1

MEAN IQs ON THE BINET AND THE WPPSI AND *t*'s BETWEEN MEANS

Test	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t</i> value*		
				WPPSI Verbal	WPPSI Performance	WPPSI Full Scale
Binet	32	95.2	13.1	5.26*	3.55*	5.57*
WPPSI Verbal	32	88.0	11.6		.34	1.06
WPPSI Performance	32	88.5	13.3			1.71
WPPSI Full Scale	32	87.1	12.8			

* *t* value on difference between means.* *p* < .01.

88.5. The Binet score was significantly higher ($p < .01$) than all WPPSI scores.

Table 2 shows the correlation coefficients between the Binet and WPPSI. All correlations for the group as a whole and also for the Negro and white subsamples are significant at the .01 level, with one exception with is significant at .05.

A comparison of the means and standard deviations of the Negro and white children is shown in Table 3. The differences between means are negligible on the WPPSI. The Binet does show a 6-point difference in favor of the white children, but this does not reach significance ($t = 1.26$). The mean and standard deviation for the Binet scores of the white children approach those expected

for the general population. The standard deviation for the Negro sample on the Binet is consistently smaller than for white Ss.

No significant differences were found between male and female Ss on the Binet or WPPSI Verbal, Performance, or Full Scale IQs. In fact, each of these scores showed less than a 1-point difference.

DISCUSSION

The Binet, in spite of the criticism found in the literature, resulted in IQ scores in the average range for the lower-class children in this study. The WPPSI scores averaged in the dull normal range. In contrast to Wechsler's results, which showed no significant differences between the Binet and the WPPSI for a group of children with a mean IQ of 90, the 8-point difference found in the present study was significant. It would appear, therefore, contrary to expectations, that

TABLE 2

CORRELATION BETWEEN THE BINET AND WPPSI

Test	Race	WPPSI Verbal	WPPSI Per- form- ance	WPPSI Full Scale
Binet	Negro	.94	.56*	.82
	White	.79	.74	.80
	Total	.81	.67	.80
WPPSI Verbal	Negro		.67	.91
	White		.80	.94
	Total		.73	.93
WPPSI Performance	Negro			.91
	White			.95
	Total			.93

Note.—The correlations, except where noted, are significant beyond the .01 level.

* $p < .05$.

TABLE 3

MEAN SCORES OF NEGRO AND WHITE CHILDREN ON THE BINET AND WPPSI

Test	Race	<i>M</i>	<i>SD</i>
Stanford-Binet IQ	Negro	92.3	9.8
	White	98.1	15.4
WPPSI Verbal IQ	Negro	88.3	10.5
	White	87.6	13.0
WPPSI Performance IQ	Negro	86.6	12.7
	White	90.5	14.1
WPPSI Full Scale IQ	Negro	86.2	11.7
	White	87.9	14.2

lower-class children earn higher IQ scores with the Binet than the WPPSI.

The correlations between the Binet and the WPPSI approximate those found with the Binet and the WISC (Schachter & Apgar, 1958), with Verbal higher than Performance. The lack of any IQ differences between races was a finding of some import, probably due to the effort to control socioeconomic factors. The lower standard deviations of the Negro sample are consistent with the findings of Kennedy et al. (1961) on the Binet ($SD = 12.4$). The homogeneity of scores probably reflects the low education level of the Negro parents which contributes to a lessening of scores in the upper range.

The lack of significant differences on the intelligence tests between male and female Ss was expected.

Subjective comparison of the WPPSI and the Binet was also part of this study. Conceptually, the examiners recognized the advantages of the Verbal and Performance division of the WPPSI. However, with respect to both administration and scoring, the examiners strongly preferred the Binet. The lower-class child is frequently shy, nonverbal, activity-oriented, and sensitive to failure. The frequent switching of materials, the briefer verbal demands, and the less apparent failure experiences made the Binet more appropriate. WPPSI disadvantages included the not infrequent changes in instructions and materials

in midtest, and the length of the testing session, which, even with the Geometric Design omitted, usually took more than 1 hour and made finishing difficult for many of the children. The examiners found several of the questions on Comprehension objectionable because of ambiguousness or emotional loadings. Having to ask for additional reasons on several of the questions made some children uncomfortable or resistive. Finally, the examiners found scoring of the WPPSI Geometric Design and Vocabulary difficult, and noted lack of scoring clarity on individual items in Comprehension. Overall, the children's enjoyment of the Binet, together with the administrative advantages, enabled the examiners to achieve better rapport and led to their willingness to accept the Binet IQ score as the more accurate.

REFERENCES

- KENNEDY, W. A., VAN DERIET, V., & WHITE, J. G. *The standardization of the Stanford-Binet Intelligence Scale on Negro elementary-school children in the Southeastern United States*. Florida: Florida State University Press, 1961.
- SCHACHTER, F., & APGAR, V. Comparison of Preschool Stanford-Binet and School Age WISC IQs. *Journal of Educational Psychology*, 1958, 49, 320-323.
- WECHSLER, D. *Manual for the Wechsler Preschool and Primary Scale of Intelligence*. New York: Psychological Corporation, 1963.

(Received September 17, 1968)

PARTICIPATION IN PSYCHOLOGICAL RESEARCH: RELATION TO BIRTH ORDER AND DEMOGRAPHIC FACTORS¹

EDWARD H. FISCHER²

AND

DAVID WINER

Connecticut Valley Hospital

Trinity College, Hartford, Connecticut

Some previous studies have indicated that research which depends on the voluntary participation of Ss may be seriously overrepresented by firstborn persons, who are thought to differ from later borns on a number of psychological dimensions. The results of four independent surveys reported here suggest that social class and religion, but not birth order, are factors in selective participation. The findings were (a) birth order had no evident effect on responding to a mailed questionnaire, either among college professors or students; (b) neither birth order nor sex predicted willingness to participate in an "extensive psychological interview"; (c) upper-class Ss were more agreeable to an interview than were lower-class Ss ($p < .05$); (d) independent of Finding c, Jewish Ss were more agreeable to an interview than were Protestants and Catholics ($p = .05$).

Rosenthal (1965) has summarized the sociological and psychological literature which describes characteristics of the volunteer S. His article served to reemphasize the point that the generality of much research is limited to those Ss willing and able to cooperate with a given study's purposes. Caution about S selection must be particularly observed where the investigator has little control over factors affecting participation and, additionally, those same factors might influence his results.

Various investigations suggest that birth order may be an important variable in the selective participation of Ss. It is contended that firstborn persons are more likely than later borns to take part in psychological studies (Capra & Dittes, 1962; Varela, 1964) and to fulfill E's requests (Wuebben, 1967). If such findings are reliable, then any study which recruits Ss on a voluntary basis may be biased by overrepresentation of firstborn persons. This is problematic because the firstborn person is thought to differ from the

later born on a variety of psychological dimensions (Sampson, 1965; Warren, 1966).

The selective cooperation of Ss became a concern for the authors in a recent study of the birth-order statuses of college professors. The result was a series of four questionnaire studies reported here.

Participation of College Professors in Survey Research

The original study was specifically planned to examine a hypothesis about birth order representation among college professors. Questionnaires requesting birth order and other information were mailed to 48 female and 48 male professors selected randomly from two college catalogues. Since only 62% of the faculty members returned the first questionnaire, it would have been impossible to generalize about their birth statuses. In light of the above-cited research, a high percentage of firstborn persons among those responding to the initial questionnaire might have meant simply that the firstborn Ss were more conscientious about completing and returning the forms.

The solution to this methodological problem was to insure a complete sample by sending a follow-up questionnaire, and then telephoning the remaining group of nonresponders. However, an analysis of the data revealed no bias in returning the forms, either by birth order or sex. Table 1 shows there was

¹ The studies were partly financed by the Faculty Research and Study Fund at Trinity College of Hartford. The data were collected with the help of Ethel Brooks and Unis Johnson at Hartford Hospital School of Nursing, Stanley L. Cohen at George Washington University, Harry L. Leonhardt at the University of Hartford, Evans Mandes of Hood College, and D. A. McCool at Rhode Island College.

² Requests for reprints should be sent to Edward H. Fischer, Department of Psychology, Connecticut Valley Hospital, Middletown, Connecticut 06457.

TABLE 1

NUMBER OF PROFESSORS WHO RESPONDED BY
QUESTIONNAIRE OR TO TELEPHONE

Response	Firstborn	Later born
Initial questionnaire	33	26
Follow-up questionnaire	11	7
Telephone	8	11

Note.— $\chi^2 = 1.53$, $df = 2$, $p < .50$.

no relationship between birth order and the manner of eliciting information from *S* ($p < .50$).

It is conceivable that the birth-order bias was absent because the *Ss*, that is, professors, differed from those where a birth-order participation effect has been found previously, that is, among students. A second study was begun, therefore, to expressly test the proposition that firstborn college *Ss* are more responsive to a mailed survey than are later-born college *Ss*.

Participation of Students in Survey Research

Ninety-six upperclassmen from a state college were sent a two-page questionnaire; ostensibly this was a follow-up to an earlier study in which birth-order and attitudinal data had been gathered. The birth-order variable had predicted interest in a college teaching career for these *Ss* (Fisher, Cohen, & de Milan, 1968).

Half each of the 56 female and 40 male students were firstborn, half were later born. The instructions urged *S* to complete and promptly return the form.

The hypothesis was that more firstborns than later borns would return the forms. In addition, one item was put in the questionnaire to test the idea that firstborn persons tend to volunteer for events involving social affiliation (Dember, 1964; Suedfeld, 1964). This item read: "I would agree to an extensive psychological interview at a later time: yes, definitely ____; maybe ____; no ____." It should be noted that this item occurred in the context of questions about teaching careers, a vocation which Fischer, Wells, and Cohen (1968) maintain has special interest for the firstborn student. Therefore the item

was expected to draw more positive responses from firstborn than from later-born students.

The results are summarized in Table 2. There was no evidence that firstborns were more responsive than later borns. In fact, more questionnaires were returned by the later borns, although this tendency was unreliable ($p < .25$).

Responses to the question about a later "extensive psychological interview" were also compared for firstborn and later-born *Ss*. While 50% of later-born *Ss* responded "yes, definitely" to the item, only 34.6% of the firstborn *Ss* did so. The difference is not significant ($\chi^2 = 1.38$, $df = 1$, $p < .25$), and the trend reverses the prediction.

Soliciting Volunteers for Research

A third survey was conducted in order to make the conditions for participation similar to the recruitment procedure used by research psychologists, and more similar to the conditions under which a relationship between birth order and volunteering has been found previously.

Recruitment forms were handed out in two undergraduate psychology classes to a total of 142 male liberal arts students. The *Ss* were told, and also read, in their instructions:

Two psychological research programs are being planned for the near future. We would like to identify persons who would be willing to take part in the studies. One of the studies will involve an extensive psychological interview dealing with personal crises faced by college-age people; the second project will involve standardization of a psychomotor (hand-eye coordination) task and the relationship of performance to personality variables. Each would take about one hour. Please indicate your interest in participation.

The students were asked to provide their names, campus addresses, and information on

TABLE 2

NUMBER OF COLLEGE STUDENTS RETURNING
QUESTIONNAIRES BY BIRTH ORDER

Response	Firstborn	Later born
Returned	26	32
Did not return	22	16

Note.— $\chi^2 = 1.57$, $df = 1$, $p < .25$.

TABLE 3

SUBJECTS' EXPRESSED WILLINGNESS TO BE
INTERVIEWED BY SOCIAL CLASS

Response	Socioeconomic class		
	I, II	III	IV, V
Definitely	49	38	41
Maybe or no	30	45	52

Note.— $\chi^2 = 6.45$, $df = 2$, $p < .05$.

birth order, family size, social class,³ and religious background. They then checked a category indicating their willingness to volunteer for one of the studies.

Analyses of the birth-order and family size data produced no significant effects. Again, slightly more later-born than firstborn Ss expressed definite interest in participation (56% and 47%, respectively, $\chi^2 = 1.15$, $df = 1$, $p > .25$). Students from large (four or more children) as compared to small (three or fewer children) families did not differ in willingness to volunteer for research ($\chi^2 = .67$, ns).

However, there were trends suggesting that both socioeconomic class and religion are related to volunteering. Over 50% of the Ss coming from Classes I, II, and III (upper- and middle-class groups) expressed definite interest in participation, while less than 30% of Ss from Classes IV and V did so. A chi-square test of the frequencies approached significance ($\chi^2 = 4.72$, $df = 2$, $p < .10$). Also, Jewish Ss appeared to be somewhat more favorable to the idea of research participation than Protestants or Catholics, although the frequency differences were not statistically reliable.

A final study, using a larger N , bore out these trends.

Demographic Variables and Volunteering for Research

The item concerning an "extensive psychological interview" was presented once more in a study of 256 nursing school and college students. The context was an anonymous

³ Social class was indexed by Hollingshead's (1957) criteria applied to the occupation and education of S's father.

TABLE 4

SUBJECTS' EXPRESSED WILLINGNESS TO BE
INTERVIEWED BY RELIGION

Response	Religion		
	Jewish	Roman Catholic	Protestant
Definitely	27	41	47
Maybe or no	13	49	53

Note.— $\chi^2 = 5.98$, $df = 2$, $p = .05$.

questionnaire which sought information about emotional problems of students. In view of the earlier research which suggests a correspondence of primogeniture and need for affiliation (Dember, 1964; Schachter, 1959), it was again expected that more firstborn than later-born Ss would be receptive to the interview. It might be said the item was loaded to attract the firstborn person. In addition to birth status, data were obtained on S's socioeconomic class and religion.

Chi-square analyses showed no effect of birth order ($\chi^2 = .01$, ns) or sex ($\chi^2 = .97$, $df = 1$, $p < .50$) on willingness to have the interview. But both the social class and religion variables were significantly related to agreement to the interview.

As can be seen in Table 3, the majority of Ss from the highest socioeconomic groups (I and II) expressed definite agreement to the interview, whereas this was not generally so for lower socioeconomic classes (IV and V). A chi-square test of the frequencies in Table 3 showed significance at the .05 level.

Table 4 shows that Jewish Ss were significantly more likely to agree to an interview than were Protestant and Roman Catholic Ss ($p = .05$). A breakdown of the Jewish group

TABLE 5

JEWISH SUBJECTS' WILLINGNESS TO BE
INTERVIEWED BY SOCIOECONOMIC CLASS

Response	Socioeconomic class		
	I, II	III	IV, V
Definitely	10	11	6
Maybe or no	7	3	3

by social class (Table 5) indicated that the finding with respect to religion was independent of the finding regarding social class. That is, Jewish Ss were fairly uniform and positive in their agreement to an interview regardless of their socioeconomic class. This was clearly not the case for Protestants and Catholics.

CONCLUSIONS

The overall results of the four independent surveys corroborate the conclusions of Ward (1964) and Wilson and Patterson (1965) that a birth-order bias ordinarily need not concern the investigator. Birth order did not appear to be a factor in responding to a mailed questionnaire, either among college professors or students. Further, there was no evident birth-order selection effect in volunteering for psychological research even when the recruitment form was slanted to attract firstborn persons (e.g., implying that the volunteer would engage in extensive interaction with a psychologist). Where slight differences did exist, it was the later born who showed somewhat more tendency to participate than the firstborn person. It may be that a significant birth-order effect occurs only when Ss are pressured to volunteer, such as in face-to-face recruitment (Capra & Dittes, 1962) or where they can gain points in a psychology course by participating in experiments (Wolf & Weiss, 1965).

The present data do indicate that at least for certain types of study, the investigator who uses volunteer Ss may be drawing disproportionately on upper-level social status groups and on Jewish Ss. The social class effect may be explainable in terms of "better" psychological-mindedness among high social status persons. The effect of religion is more difficult to explain, but it may be related to Rotter's (1967) finding that Jewish students express more interpersonal trust than Protestants and Catholics. In the current academic setting where "psychological research" may have many dubious connotations, a funda-

mental trust of professionals and existing institutions may be requisite to volunteering for research.

REFERENCES

- CAPRA, P. C., & DITTES, J. E. Birth order as a selective factor among volunteer subjects. *Journal of Abnormal and Social Psychology*, 1962, 64, 302.
- DEMBER, W. N. Birth order and need for affiliation. *Journal of Abnormal and Social Psychology*, 1964, 68, 555-557.
- FISCHER, E. H., COHEN, S. L., & DE MILAN, J. Ordinal position and interest in the academic career. Paper presented at the meeting of the Eastern Psychological Association, Washington, D. C., April 1968.
- FISCHER, E. H., WELLS, C. F., & COHEN, S. L. Birth order and expressed interest in becoming a college professor. *Journal of Counseling Psychology*, 1968, 15, 111-116.
- HOLLINGSHEAD, A. B. Two factor index of social position. Unpublished manuscript, Yale University, 1957.
- ROSENTHAL, R. The volunteer subject. *Human Relations*, 1965, 18, 389-406.
- ROTTER, J. B. A new scale for the measurement of interpersonal trust. *Journal of Personality*, 1967, 35, 651-665.
- SAMPSON, E. E. The study of ordinal position: Antecedents and outcomes. In B. Maher (Ed.), *Progress in experimental personality research*. New York: Academic Press, 1965.
- SCHACHTER, S. *The psychology of affiliation*. Stanford: Stanford University Press, 1959.
- SUEDFELD, P. Birth order of volunteers for sensory deprivation. *Journal of Abnormal and Social Psychology*, 1964, 68, 195-196.
- VARELA, J. A. A cross-cultural replication of an experiment involving birth order. *Journal of Abnormal and Social Psychology*, 1964, 69, 456-457.
- WARD, C. D. A further examination of birth order as a selective factor among volunteer subjects. *Journal of Abnormal and Social Psychology*, 1964, 69, 311-313.
- WARREN, J. R. Birth order and social behavior. *Psychological Bulletin*, 1966, 65, 38-49.
- WILSON, P. R., & PATTERSON, J. Sex differences in volunteering behavior. *Psychological Reports*, 1965, 16, 976.
- WOLF, A., & WEISS, J. H. Birth order, recruitment conditions, and volunteering preference. *Journal of Personality and Social Psychology*, 1965, 2, 269-273.
- WUEBBEN, P. L. Honesty of subjects and birth order. *Journal of Personality and Social Psychology*, 1967, 5, 350-352.

(Received September 18, 1968)

DIRECTION OF AFFECTIVE INFLUENCE IN PSYCHOTHERAPY¹

KENNETH I. HOWARD²

Northwestern University

MERTON S. KRAUSE

Institute for Juvenile Research, Chicago

AND DAVID E. ORLINSKY³

University of Chicago

Each participant in 45 patient-therapist pairs completed a structured response questionnaire after each of a series of individual psychotherapy sessions. In these questionnaires they reported on the feelings experienced by each during the session. A method of causal analysis of nonexperimental data was applied to these reports. These analyses suggested that for relatively inexperienced therapists, the patients' positive feelings were a sufficient condition for the therapists' positive feelings. The reverse was the case for more experienced therapists. For negative feelings, however, therapists' feelings of uncertainty and apprehension were generally a sufficient condition for the patient to experience dysphoric affect. Extremely positive or negative affective experiences of either participant tended to preclude and be precluded by opposite experiences in the other.

The feelings that patients and therapists experience during a psychotherapeutic session are events of central importance to the treatment process. Exploration of affect is a prime task of the patient and affective cues are used by the therapist in understanding and relating to the patient. Even aside from these functions, of course, the psychotherapy relationship is usually affectively alive for the two participants.

In a previous paper evidence was presented which indicated that patient and therapist share the dominant mood or affective tone of a psychotherapeutic session, that is, there tends to be a commonality of affective experience (Howard, Orlinsky, & Hill, in press; Lennard & Bernstein, 1960). It does not seem reasonable, however, that common affective experience merely arises independently or spontaneously for both participants in psy-

chotherapy. Most likely, for a particular pair in a particular session, either the patient or the therapist affectively influences the other, so that the feelings of one are responded to affectively by the other.

The research reported in this paper is aimed at providing a better understanding of affective processes in individual psychotherapy. More specifically, it is concerned with the direction of affective influence. Two distinct and, in effect, contradictory positions have been reported in the literature. Fiedler (1953), Truax (1963), Wilson and Hannon (1968), Rausch and Bordin (1957), and Patterson (1967) all stress the prepotence of the therapist's feelings. Their research seems to indicate that the therapist sets the affective tone of the relationship. On the other hand, Heller, Myers, and Kline (1963), Bohn (1965), Snyder (1963), and Moos and Clemes (1967) all stress the reactivity of the therapist's feelings. Their research seems to indicate that the patient sets the affective tone and the therapist reacts to the patient's feelings. This paper, using an adaptation of a method discussed by Blalock (1964), investigates the evidential support for, or contradiction of, these two positions.

The ideal way to carry out such an investigation would be to conduct a controlled ex-

¹ An extended version of this paper was presented at the annual meeting of the Society of Multivariate Experimental Psychology, Austin, November 1968. This study was supported in part by General Research Support Grant FR-05666-01 from the National Institutes of Health to the Institute for Juvenile Research.

² Also at the Institute for Juvenile Research. Requests for reprints should be sent to Kenneth I. Howard, Institute for Juvenile Research, 232 East Ohio Street, Chicago, Illinois 60611.

³ Also at the Institute for Juvenile Research.

periment in which one participant's affect were somehow manipulated and the other's affective responses were assessed. The probative power of this experiment, perfectly performed, derives from the control over the time sequence of changes in the independent and dependent variables and the stochastic minimization of external confounding (Krause, 1967). Experiments are basically intended to support or contradict propositions of the form, "A is a sufficient condition for B," that is, whenever A occurs B will occur, although B may occur in the absence of A. Without control of the time sequence, the evidence for A as a sufficient condition for B cannot be distinguished from the evidence for B as a necessary condition for A. This latter proposition means that A will only occur when B has occurred, although B may occur at other times. If the time sequence between A and B cannot be specified, the choice between these different causal propositions must be made on other evidence, for example, plausibility. This is precisely what happens in the case of causal analyses of non-experimental data (Blalock, 1964).

It should be noted, however, that several causal propositions are contradicted by any data which support (equally) the two mentioned propositions. We can construe the data as consisting of a simple count of the joint occurrence of the various combinations of A, not-A, B, and not-B (a fourfold table). If we find that cells A,B, not-A,B, and not-A, not-B are occupied and the cell A,not-B is not occupied, these results would support the two propositions mentioned in the previous paragraph (see Figure 1). Figure 1 shows that B occurs in the presence or absence of A, but that A occurs only in the presence of B. A is thus a sufficient but not necessary condition of B, or B is a necessary but not sufficient condition of A. However, the following additional propositions are contradicted by these results: (a) A is a necessary and sufficient condition for B; (b) B is a necessary and sufficient condition for A; (c) B is a sufficient condition for A; and (d) A is a necessary condition for B. Thus, causal analysis of nonexperimental data can at least *limit* the alternative explanations for that set of data (Blalock, 1964; Krause, 1967). It also

	not-B	B
not-A	X	X
A	O	X

FIG. 1. Evidence consistent with the proposition that A is a sufficient condition for B.

represents a more complete use of correlational data.

As noted above, experiments generally are concerned with documenting the *sufficiency* of an independent variable, A, not-A, for producing a certain set of states of the dependent variable, B, not-B. The corresponding *necessity* proposition, that is, that B is necessary for A, is ruled out by the physical control imposed. For nonexperimental data the choice between such corresponding propositions of necessity and sufficiency is generally based on plausibility. Regarding the analysis of affective influence, the choice was based on the following considerations. It seems implausible that the therapist's mood is a *necessary* condition for some affective state of his patient, for example, that the patient can never be uncommonly sad unless the therapist is uncommonly sad. The *necessity* of certain patient moods for therapist affective experience seems even more implausible, for example, that the therapist can never feel uncommonly good unless the patient feels uncommonly good. These notions of implausibility derive both from clinical experience and knowledge of the volatility and multiple causation of affective states.

In the light of the preceding considerations, the most plausible rendering of the two positions (hypotheses) concerning direction of affective influence in psychotherapy would be: (a) therapist feelings are a sufficient condition for patient feelings; and (b) patient feelings are a sufficient condition for therapist feelings.⁴ An examination of the empirical

⁴ Data that are consistent with Hypothesis a are also consistent with the proposition that (c) the absence of a certain therapist feeling is sufficient for the absence of the corresponding patient feeling. Data

support (based on the data of the Psychotherapy Session Project) for each of these hypotheses is the major focus of this paper.

METHOD

As part of a larger study, a panel of patients and therapists completed structured-response questionnaires immediately after each of a series of psychotherapy sessions. The questionnaires (Therapy Session Report), methods of data collection, and samples of therapists and patients have been described elsewhere (Howard et al., in press; Howard, Orlinsky, & Hill, 1969; Orlinsky & Howard, 1966; Orlinsky, Howard, & Hill, 1969).

Sample

The sample employed for the analyses reported in this paper consisted of 45 patient-therapist pairs. The members of each pair completed Therapy Session Reports (TSRs) after from 5 to 58 psychotherapeutic sessions, creating a pool of 893 sessions for which there were joint reports.

The 45 patients were female and most were single and employed. They ranged in age from 19 to 60, but the majority were under 30. Fifteen percent were nonwhite, 88% had at least a high school education (65% had attended college), and 59% had had some previous psychotherapy. All patients were being treated in once-weekly individual psychotherapy, each session scheduled for 45 minutes.

Of the 15 therapists in this sample 13 were male. The therapists had been trained in psychiatry, clinical psychology, or psychiatric social work; their median experience level was 6 years. Over half of the therapists had had personal therapy. Their theoretical orientations were generally "dynamic-eclectic" in that no therapist represented a single school or theory. Most therapists reported on sessions for from 1 to 3 patients, but two therapists reported on as many as 15 patients.

Instruments

There are two parallel forms of the TSR, one for patients and the other for therapists. Within each form are two sections which survey the feelings experienced by the participants during the session. An early section asks each to report how the patient felt, and provides a list of 33 feelings. A later section asks each to report how the therapist felt and provides a list of 25 feelings. For each feeling the respondent is instructed to circle a number (0 = No; 1 = Some; 2 = A Lot) to indicate the intensity of its presence during the session.

In a previous paper (Howard et al., in press), the results of factor analyses of these feeling lists were

consistent with Hypothesis *b* are also consistent with (*d*) the absence of the corresponding therapist feeling. Both *c* and *d* seem quite implausible with regard to feelings like happiness and sadness, which shall be discussed.

reported. Although a more highly differentiated structure could be defined, there was considerable support for a two-factor structure of each list. The two factors closely resembled the positive affect and negative affect dimensions reported by Snyder (1961). Patients' positive affective experience (P+) was essentially defined by the following feelings: grateful, relieved, hopeful, pleased, relaxed, confident, affectionate, and sexually attracted. Therapists' positive affective experience (T+) was mainly constituted by optimistic, cheerful, playful, effective, pleased, sympathetic, close, and affectionate. Patients' negative affective experience (P-) was essentially defined by angry, impatient, cautious, inhibited, withdrawn, frustrated, tearful, depressed, inadequate, anxious, guilty, strange, and embarrassed. The feelings which comprised therapists' negative affective experience (T-) were apprehensive, perplexed, unsure, and demanding.

For each session reported on by each participant, four affect scores were computed: patients' positive and negative affect (P+ and P-) and therapists' positive and negative affect (T+ and T-). These scores were simply the sums of the participant's responses to the feeling words in each composite, and represent the data base for statistical analyses.

Three pairings of judgments of positive or negative affective process could be derived. *Patient-perceived* positive affective process was embodied in her reports of P+ and T+. *Therapist-perceived* positive affective process could be derived from his reports of P+ and T+. Finally, "actual" positive affective process could be obtained from the patient's reports of P+ and the therapist's reports of T+. The same three pairings were possible for negative affect. The analytic model for all comparisons, however, was the same.

The major purpose of the analyses was to discover which, if either, of the two causal propositions regarding direction of affective influence was consistent with the empirical data. The same procedure was followed for each of the six possible comparisons outlined above. For purposes of exposition "actual" positive affective process will be used to illustrate the method of analysis. The analyses were performed for each patient-therapist pair, separately, over their common sessions.

The basic idea was to construct a fourfold table consisting of the frequency of joint occurrence of the four conditions: P+, T+; not-P+, not-T+; P+, not-T+; and not-P+, T+. If both off-diagonal cells (not-P+, T+ and P+, not-T+) were empty, this would be consistent with both hypotheses. If only not-P+, T+ were empty, this would be consistent only with the therapist influence (TI) interpretation, since the therapist's feeling could occur without the patient's but not vice versa. Finally, if P+, not-T+ were the only empty cell, only the patient influence (PI) hypothesis would be consistent with the data, since the therapist would only feel good when the patient felt good.

The data, however, were not dichotomized in terms of the presence or absence of T+ or P+; the

scores were continuous. Consequently, some criterion for dichotomization was necessary in order to create the basic fourfold tables required by the analytic model. Since the purpose of this analysis was to differentiate between causal propositions about the direction of affective influence, a criterion was needed that would maximize the chances of distinguishing cases of therapist predominance from cases of patient predominance. To this end, the criterion of "the largest empty off-diagonal quadrant" was devised. The implementation of this criterion was as follows: (a) For each patient-therapist pair, T+ and P+ scores were cross-tabulated over their common sessions.⁶ (b) These scatter plots were inspected to determine the largest off-diagonal area, including a corner, which had no entries. (c) A fourfold table was constructed using the cutoff points indicated by Step b. Finally, (d) the fourfold table was inspected to determine which of the causal propositions was consistent with the results.

The reader should note that this method of dichotomizing the data is based on the assumption of a certain underlying model for data generation. (a) The data for each patient-therapist pair are assumed to be an unbiased sample from a longer finite series of their interviews, that is, trends in direction of influence are precluded, but (b) each datum point is assumed to be free of measurement error. Thus, we assume a population distribution in cluster shape rather than (as has become conventional) a population line with measurement error. In the method of dichotomization, two points can determine the frame for defining quadrants, and single points can determine what shall be the largest empty quadrant. Under Assumptions a and b, these characteristics of the method are quite reasonable. Finally, (c) it is assumed that one of the two opposite hypotheses is true: either the therapist's feeling is a sufficient condition for the patient's or vice versa. These assumptions then imply that whatever dichotomization yields the best discrimination between the two hypotheses, in c, is the best estimate of the true or latent dichotomization, though it is subject to sampling error. (Note that there is no assumption here of the same hypothesis being true for all patient-therapist pairs.)

RESULTS

The above procedures were carried out on the data from each of the 45 patient-therapist pairs for the three comparisons involving positive affect and the three involving negative affect. In this section the focus is on the implications of these results for therapy process in general. The next section explores

the differences in direction of influence associated with patient and therapist characteristics.

Positive Affect

Comparisons of the therapists' reports of positive feelings with the patients' reports of positive feelings did not reveal predominant support for either of the hypothesized causal relationships. Seventeen pairs were consistent with the therapist influence (TI) hypothesis, 19 pairs were consistent with the patient influence (PI) hypothesis, and 9 pairs equally supported both. The perceived feeling process of the patients (their reports of their own positive feelings compared to their reports of their therapists' positive feelings) yielded essentially the same results: 11 TI, 19 PI, and 15 indeterminate. The perceived feeling process of therapists (their reports of their own positive feelings compared with their reports of their patients' positive feelings) was more consistent with the patient influence hypothesis: 9 TI, 23 PI, and 13 indeterminate. Therapists, thus, tended to see patients' positive feelings in a manner consistent with their being sufficient for the therapists' own experiences of positive feelings. But the convergence of the three perspectives was weak, indicating that no general statement could be sustained for patient-therapist pairs.

Negative Affect

The comparisons involving negative (dysphoric) feelings revealed a more consistent trend. Analysis of patients' reports of negative feelings and therapists' reports of negative feelings supported the therapist influence hypothesis: 26 therapist influence, 9 patient influence, and 10 indeterminate. The results obtained from the comparisons involving patient-perceived negative feeling process were: 27 TI, 10 PI, and 8 indeterminate. Therapist-perceived negative feeling process was also consistent with these results: 29 TI, 10 PI, and 6 indeterminate. Binomial tests of the splits between patient influence and therapist influence were statistically significant for all three comparisons, beyond the .01 level. For negative affect, then, it appeared that the presence of negative affective experiences for the therapist may usually be sufficient to pro-

⁶ The authors most gratefully acknowledge the diligent assistance of Craig Barton, Len Sushinsky, Jim Trattner, and Howard Ullan in performing this part of the data analysis.

vide a negative affective experience for the patient; that is, although the patient may feel unusually bad in any particular session, she is considerably more likely to feel bad if the therapist does.

The degree of case-by-case agreement between the three perspectives, patient's (P), therapist's (T), and "actual" (A), was 38% (PT), 36% (PA), and 42% (TA) for positive feelings and 58% (PT), 53% (PA), and 56% (TA) for negative feelings. None of the six 3×3 tables from which these values were computed had a significant chi-square of association at the .01 level. Thus, the three perspectives may not be considered convergent means of assessing the same facts of influence. Because the "actual" perspective compares the results of two independent assessments, each by the person who has the best information about how *he* feels, we are inclined to treat that data as most valid for this study.

Although the results differed for positive and negative feelings, direct comparisons revealed an interesting interdependency. When fourfold tables were constructed comparing patient negative feelings with therapist positive feelings (P- vs. T+) and patient positive feelings with therapist negative feelings (P+ vs. T-), there was strong evidence that the experience of a marked degree of one type of feeling (e.g., positive affect) for one participant precluded the experience of an extreme degree of the other type of feeling (e.g., negative affect) by the other participant. If one felt especially bad, the other tended not to feel especially good; if one felt especially good, the other tended not to feel especially bad. Unfortunately, both the necessity and the sufficiency interpretations of these findings are equally plausible, that is, the following two statements seem equally defensible: (a) the therapist cannot feel especially effective unless the patient does not feel especially bad; (b) the patient will not feel especially bad when the therapist feels especially effective. However, the results are inconsistent with propositions that (c) the absence of a particular extreme affective experience for one participant is sufficient to elicit the presence of the opposite extreme affective experience for the other, and (d) the presence of a par-

ticular extreme affective experience for one participant is necessary for the absence of the opposite extreme affective experience of the other.

Results of Individual Differences Analyses

Clearly there were individual differences as to which participant was affectively predominant. The purpose of the analysis in this section was to determine what individual characteristics of patients or therapists were differentially associated with patient or therapist affective predominance.

In all, 10 patient characteristics were explored: age, education, marital status, religion, family size, birth order, parental separation, age at familial disruption, diagnosis, and previous psychotherapeutic treatment. The 11 therapist characteristics investigated were: age, religion, social class of origin, marital status, parental status, sex, family size, birth order, profession, experience, and personal psychotherapy. Chi-square was employed to screen these 21 personal characteristics for significant associations with the various patterns of affective influence. Since the purpose of this analysis was exploratory, the .10 significance level was employed in screening. Only entries which contributed a disproportionately large amount to a significant chi-square will be discussed.

No patient characteristic was significantly differentially associated with direction of influence in either affective process. No therapist characteristic was differentially associated with direction of influence in negative feeling process, but three therapist characteristics were related to positive process. These were (a) experience ($p < .01$), (b) age ($p < .10$), and (c) birth order ($p < .10$). For those pairs where the evidence was consistent with the proposition that the patient's positive feelings were sufficient for the therapist to experience unusually positive feelings (PI), 74% of the therapists had had less than 6 years' experience. For the TI pairs, on the other hand, 76% had had 6 or more years' experience in the practice of psychotherapy. Seventy-nine percent of the therapists in PI pairs were under 35 years of age, while 69% of the TI therapists were 35 or older. Finally,

only 11% of the PI therapists had at least one older same-sex sibling, while 41% of the TI therapists had at least one older same-sex sibling in their family of origin. On the whole the 21 independent variables available in this study were not strongly predictive of affective predominance, with the exception of therapist experience.

DISCUSSION

When a patient and her therapist indicate how they felt in each immediately preceding therapy session, they give evidence bearing on which of them is the more affectively influential. It was found that for most of the patient-therapist pairs, the therapist was the more influential of negative feelings. This seems quite understandable, because the patient, but not the therapist, is supposed to be feeling badly. The patient's bad feelings are part of his task and are no reason for a competent therapist to actually feel badly himself, certainly no reason for feeling inadequate as a therapist. However, if he does feel this way, that is good reason for the patient to feel negative in and about her therapy. She has presumably come preoccupied with her problem and expecting help from an expert. If he feels badly, that is, inadequate, then how can she hope to get help from him? If he feels good, that may be reassuring but not necessarily enough to raise her spirits.

A second finding, that for less experienced (and younger) therapists, patients' positive feelings are more likely to be sufficient for them to feel good, also seems quite understandable. A less experienced therapist is more likely to be responsive to his patients' positive feelings, because he needs evidence of his competence and of a good prognosis. Patients' positive feelings are such evidence. More experienced therapists probably tend to have a more secure concept of themselves as therapists, and so are less affectively responsive to their patients in this way. It may also be that being more secure, they can convey a more clear and certain sense of their extreme positive affect than can the less experienced therapists.

How a therapist feels about his treatment of a patient and the effect of these feelings on

the patient may be manipulable factors. If therapists could be made to feel unconfident and unsuccessful, for example, through critical supervision, stronger evidence of the sufficiency of their negative feelings for patient negative feelings could be had. Similarly, if patients could be made to feel more positively, for example, through an assessment interview, the second finding could be strengthened. Without such manipulations, of course, we must rest on the more inferential evidence available from the sort of retrospective data of this study or the somewhat better evidence derivable from process analysis of single therapy sessions. In the meantime, it would seem advisable for therapists to experiment with various ways of handling their dysphoric sessions. Perhaps canceling such sessions, clarifying the patient's perception of her therapist's mood, taking special notice of the affective interaction, or consciously managing one's affective display would yield better results on such hours. Surely the treatment must be adapted to the therapist as well as to his patient.

REFERENCES

- BLALOCK, H. M. *Causal inferences in nonexperimental research*. Chapel Hill: University of North Carolina Press, 1964.
- BOHN, M. J., JR. Counselor behavior as a function of counselor dominance, counselor experience and client type. *Journal of Counseling Psychology*, 1965, 12, 346-352.
- FIEDLER, F. E. Quantitative studies on the role of therapists' feelings toward their patients. In O. H. Mowrer (Ed.), *Psychotherapy theory and research*. New York: Ronald Press, 1953.
- HELLER, K., MYERS, R. A., & KLINE, L. V. Interviewer behavior as a function of standardized roles. *Journal of Consulting Psychology*, 1963, 27, 117-122.
- HOWARD, K. I., ORLINSKY, D. E., & HILL, J. A. Affective experience in psychotherapy. *Journal of Abnormal Psychology*, in press.
- HOWARD, K. I., ORLINSKY, D. E., & HILL, J. A. Content of dialogue in psychotherapy. *Journal of Counseling Psychology*, 1969, 16, 396-404.
- KRAUSE, M. S. Proving causal propositions: The foundations of program and experiment design. *Multivariate Behavioral Research*, 1967, 2, 349-376.
- LENNARD, H. L., & BERNSTEIN, A. *The anatomy of psychotherapy*. New York: Columbia University Press, 1960.

- MOOS, R., & CLEMES, S. R. Multivariate study of the patient-therapist system. *Journal of Consulting Psychology*, 1967, 31, 119-130.
- ORLINSKY, D. E., & HOWARD, K. I. *Therapy session report, Forms P and T*. Chicago: Institute for Juvenile Research, 1966.
- ORLINSKY, D. E., HOWARD, K. I., & HILL, J. A. The patient's concerns in psychotherapy. *Journal of Clinical Psychology*, 1969, in press.
- PATTERSON, C. Divergence and convergence in psychotherapy. *American Journal of Psychotherapy*, 1967, 21, 4-17.
- RAUSCH, H. L., & BORDIN, E. S. Warmth in personality development and in psychotherapy. *Psychiatry*, 1957, 20, 351-363.
- SNYDER, W. U. *The psychotherapy relationship*. New York: Macmillan, 1961.
- SNYDER, W. U. *Dependency in psychotherapy*. New York: Macmillan, 1963.
- TRUAX, C. B. Effective ingredients in psychotherapy: An approach to unraveling the patient-therapist interaction. *Journal of Counseling Psychology*, 1963, 10, 256-263.
- WILSON, G. T., & HANNON, A. E. Behavior therapy and the therapist-patient relationship. *Journal of Consulting and Clinical Psychology*, 1968, 32, 103-109.

(Received September 20, 1968)

(Continued from page 589)

- Internal-External Locus of Control and National Stereotypes in Denmark and the United States: Oscar A. Parsons,* John M. Schneider, and Arne S. Hansen: University of Oklahoma Medical Center, 800 Northeast Thirteenth Street, Oklahoma City, Oklahoma 73104.
- Perception of Disability by the Nondisabled: A. P. MacDonald, Jr.* and Janet Hutchinson: Research and Training Center, Old Mountainlair, Room 41, West Virginia University, Morgantown, West Virginia 26506.
- An Evaluation of the Figure Drawing Technique: Reliability, Factorial Structure, and Diagnostic Usefulness: Peter T. Adler*: Department of Psychiatry, New York Medical College—Metropolitan Hospital, 1901 First Avenue, New York, New York 10029.
- Verbal Conditioning and Therapeutic Change: Paul S. Lapuc and Morton Harmatz*: Department of Psychology, University of Massachusetts, Amherst, Massachusetts 01002.
- Effect of Accurate Forecasts on Length of Hospital Stay of Psychiatric Patients: Benjamin F. McNeal,* Roy Johnston, and Victor A. Aspromonte: Veterans Administration Hospital, Canandaigua, New York 14424.

* Asterisk indicates for whom address is supplied.

PREDICTION ENHANCEMENT THROUGH THE USE OF MODERATOR VARIABLES

JAMES E. COWDEN¹

Wisconsin Department of Health and Social Services

In order to increase the ability of clinicians to predict the subsequent behavior of inmates in a juvenile correctional institution, moderator variables (i.e., variables which were correlated with degree of predictability of inmates as measured by the relative magnitudes of the difference (D) scores between predictor and criterion variables) were developed using a construction sample of 76 Ss. These moderator variables were then used to remove less predictable Ss from a cross-validation sample, resulting in a moderate increase in the correlation between predictor and criterion variables when one-third of the least predictable Ss were removed, and a more substantial increase when one-half of the least predictable Ss were similarly removed. Hence, this method showed considerable potential value as a means of facilitating more accurate classifications of inmates.

The problem of how to increase the ability to predict a given criterion is traditionally resolved by increasing the reliability and validity of the predictor and criterion variables, by more precise measurements of these variables, or by increasing the number of predictors. In most cases, the use of these techniques insures at least a moderate level of prediction. However, when not more than two or three predictors of moderate reliability and validity are available, and it is important to achieve the maximal possible degree of prediction, other means of enhancing prediction must be found. Ghiselli (1960, 1963) has concerned himself with this problem, with the goal of predicting as accurately as possible which employees would subsequently show the best job performance. He pointed out that except in the cases where correlations between predictor and criterion variables are quite high, there is typically a significant variation between Ss in difference (D) scores between predictor and criterion scores. For any given set of predictor and criterion variables, individuals could thus be differentiated along a dimension of predictability, as measured by the magnitudes of these D scores. If other tests or ratings were then given to the same Ss and found to be correlated with these D

scores, these ratings or tests could be used to select subsequent samples of Ss (particularly if similar to the original sample) having relatively small D scores. Using this technique, Ghiselli (1956) was able to select certain Ss from a cross-validation sample who were presumed to be more predictable than average. In support of his hypothesis, he found that the correlation between predictor and criterion variables was significantly higher for this subsample than for the remaining Ss in the sample.

In the present study, an attempt was made to use the above techniques in somewhat modified form to assess their efficacy in enhancing the prediction of institutional and postrelease adjustment of delinquent boys. Clinicians in correctional institutions are typically asked to classify and make recommendations about newly admitted inmates based on judgments as to their probable subsequent adjustment within the institution. The results of previous studies conducted at two Wisconsin correctional institutions for delinquent boys suggested that age at first admission to the institution and global prognostic ratings by clinicians and social workers were significantly related to institutional and postrelease adjustment (Cowden, 1966, Cowden & Pacht, 1967). While these two variables were among the best predictors of the subsequent adjustment of delinquent boys, the magnitudes of the correlations between the individual pre-

¹ Requests for reprints should be sent to James E. Cowden, Department of Health and Social Services, Division of Corrections, 1 West Wilson Street, Madison, Wisconsin 53701.

dictor and criterion measures were not very high (within the range of .26 to .54). Adding other predictors (using a multivariate approach) increased the multiple correlation coefficient significantly, to about .68. However, this approach did not appear very feasible mainly because the process of assigning weights to 6-10 predictors and using them in combined form to predict adjustment was found to be too complicated for routine classification purposes. Second, the multiple correlation coefficients tended to show a significant decline when the multiple-regression equations were used on new (cross-validation) samples, most likely because the use of multiple predictors allowed a variety of sources of error to operate. These considerations suggested that more attention might profitably be focused on methods of increasing the predictive value of individual predictors, as opposed to using a multiple regression approach. Hence, an attempt was made in this study to enhance the prediction of institutional and postrelease adjustment through selecting more predictable Ss.

METHOD

Minnesota Counseling Inventory (MCI) scale scores (Berdie, 1957) and ratings of personality, social, and demographic variables derived from case folders of 152 consecutive first admissions to the Wisconsin School for Boys at Wales, Wisconsin, were used as sources of data. This sample was typical of the population of boys institutionalized as delinquents at juvenile correctional institutions in Wisconsin (as well as in juvenile correctional institutions in most other states). All boys in Wisconsin between the ages of 12-18 who are committed as juvenile delinquents to the State Department of Health and Social Services must be received at the Reception Center of the Wisconsin School for Boys at Wales, Wisconsin (the institution from which the sample was drawn). Approximately half of the boys are then retained at the School for Boys at Wales and half are transferred to a similar juvenile correctional institution, the Kettle Moraine Boys School at Plymouth, Wisconsin. The boys at these two institutions are very similar in most respects, though more seriously disturbed boys tend to be retained at Wales. Statistics compiled by the Bureau of Research (June 1967 and September 1967) reveal the following characteristics of the boys in this population during 1966. The median age of boys at the time of their first admission was 15.9, they had completed 8.8 grades in school, their achievement grade level (California Achievement tests) was slightly below the seventh grade level, and 34.2% were nonwhite. The major

offenses for which they had been committed included auto theft, (28%), burglary (23%), truancy (11%), uncontrollability (10%), theft (9%), aggravated assault (5%), and sex misconduct (4%). Sixty-eight percent had been involved with one or more companions in their delinquent behavior, while 32% had no such companions. Their mean length of stay within the institution in 1966 was 4.5 months and their median age at release was 16 years, 3 months. Fifty-one percent of those in the sample came from homes in which both parents were married, while the remainder came from homes in which the parents were separated, divorced, not married, or widowed. Seventy percent of the boys came from lower socioeconomic class levels, while 30% came from middle-class backgrounds. Fifty-five percent of those released in 1966 were released for the first time from a juvenile correctional institution.

The sample assessed in this study appeared to be representative of the above population of delinquent boys in the characteristics described. Seventy-six Ss of the 152 in the sample were then randomly assigned to the construction sample. Thirty-seven predictor variables and several criterion measures of institutional and postrelease adjustment were then intercorrelated. Three predictor variables showing at least a moderate relationship with one or more of the criterion measures were then selected. The distributions of these predictor variables and the corresponding criterion variables were then transformed into standard (Z) distributions with means of 50 and standard deviations of 10. The predictor and criterion score distributions were thus equally weighted in constructing D-score distribution, constructed by subtracting the criterion Z scores (Z_c) of each S from his predictor Z scores (Z_p). The signs of the D scores were disregarded in deriving mean D scores, since either positive or negative D scores constituted errors in prediction.

The data available for each S then included three D scores in addition to his 37 predictor ratings and three criterion ratings. These three D scores were then correlated with the 37 predictor variables in order to select the three predictor variables which correlated most highly with the D scores, that is, which of these predictor variables functioned most effectively as moderators.

Three variables were selected in this manner to function as moderators for each of the three predictor-criterion combinations. These three moderator variables were then used in turn to subdivide Ss in the cross-validation sample into subgroups showing different degrees of predictability. Correlation coefficients as well as mean D scores were then calculated between the three predictor and criterion variables in turn for the following: the entire cross-validation sample; the two-thirds of the cross-validation sample which was most predictable, according to their scores on the moderator variable used; and finally, the one-half of the cross-validation sample which was most predictable according to their scores on the moderator variable used. This process was repeated for each of the three predictor-criterion

TABLE 1

CORRELATIONS BETWEEN PREDICTOR AND CRITERION VARIABLES FOR ALL SUBJECTS
AND FOR SUBJECTS SELECTED THROUGH THE USE OF MODERATOR VARIABLES

Moderator variables	Predictor variables	Criterion variables	Construction sample (<i>N</i> = 76)	Cross-validation sample		
				<i>N</i> = 76	<i>N</i> = 54	<i>N</i> = 38
MCI V scale	COPR	SWFR	.54	.47 $\bar{X}_{Diff} = 7.99$.69 $\bar{X}_{Diff} = 8.06$.73 $\bar{X}_{Diff} = 7.45$
MCI V scale	SWPR	SWFR	.47	.41 $\bar{X}_{Diff} = 8.03$.53 $\bar{X}_{Diff} = 7.21$.67 $\bar{X}_{Diff} = 5.91$
COPR	CLPR	PAPR	.44	.44 $\bar{X}_{Diff} = 8.80$.46 $\bar{X}_{Diff} = 8.52$.83 $\bar{X}_{Diff} = 4.47$

Note.—COPR = counselor prognostic ratings; SWPR = social workers' prognostic ratings; CLPR = clinicians' prognostic ratings; SWFR = social workers' final ratings of institutional adjustment; PAPR = parole agents' ratings of postrelease adjustment.

distributions assessed, allowing an evaluation to be made of the degree to which each of the moderator variables increased the predictive capabilities of each of the three predictor variables.

RESULTS AND DISCUSSION

The three predictor variables selected for use included prognostic ratings by counselors, social workers, and clinicians (psychologists or psychiatrists). These were essentially predictions, on 7-point scales, of the adequacy of the boys' subsequent institutional and post-release adjustment, and were derived from reports written shortly after each boys' commitment to the institution. They were found to be moderately reliable, with reliability coefficients of .62, .68, and .76. The criterion variable associated with the first two predictor variables was the final rating of institutional adjustment (derived from social workers' reports submitted just prior to the boys' release from the institution). In the third case, the criterion variable consisted of ratings of post-release adjustment, derived from probation-parole agents' reports of the boys' postrelease adjustment for the first 3 months after being released from the institution. Both of these criterion ratings could be rated reliably, as shown by previously obtained reliability coefficients of .74 and .70, respectively. The V scale of the Minnesota Counseling Inventory (on which higher scores indicate greater defensiveness) functioned as a moderator variable in the first two cases, that is, higher scores on the V scale had been found (on the construction sample) to be correlated significantly

with lower D scores (greater predictability). Counselors' global prognostic ratings functioned in the third case as a moderator variable, with higher scores (more positive prognostic ratings) correlated with lower D scores (greater predictability) on the construction sample. Hence, these ratings were thus used in the third case to selectively remove the less predictable boys from the cross-validation sample.

As shown in Table 1, when the moderator variables were used to remove one-third of the least predictable boys from the cross-validation sample, the correlations between predictor and criterion variables showed a slight to moderate increase (.22, .12, and .02) in the three cases. When half of the least predictable boys were similarly removed, the correlation between predictor and criterion variables showed a much more consistent and marked increase (.26, .26, and .39) in the three cases.

These results suggest that selectively removing less predictable Ss through the use of moderator variables markedly enhances the clinician's ability to predict their subsequent behavior, particularly when half of the Ss are removed in this manner. One obvious limitation of this approach is that it excludes the one-third or one-half of the Ss who are relatively unpredictable. Although relatively unpredictable, these inmates must still be classified by the clinician using these moderators. Clinicians could, of course, use this information as a rationale for being more tentative in classifying these inmates. On the other hand,

if it is important to make accurate predictions concerning this residual group of relatively less predictable Ss, additional moderator variables could be used in similar fashion to enhance predictions concerning their subsequent behavior. Moderator variables appear to be significantly correlated only with certain predictor-criterion difference distributions. Hence, if certain Ss are found to be unpredictable using one moderator, this does not mean that they will be similarly unpredictable using another. Using two moderator variables successively is, of course, a more complex and time-consuming procedure than using only one. However, this is less complicated than using a number of predictor variables, as would be done using a multiple-regression approach.

Although some problems thus exist in the use of moderator variables, this technique appears potentially more useful under certain conditions than a multiple-regression approach, not only because the resultant correlation coefficients are significantly higher than those achieved through the use of multiple-regression techniques, but also because the method is simpler to use by institutional staff members in classifying newly admitted inmates. Although the present study involved the selection of moderators simply on the basis of the magnitudes of their correlations with difference scores, it is possible that as more information becomes available concerning their stability and generality, a

theoretical rationale could be developed which would facilitate their selection and use. Subsequent studies will be undertaken with this goal in mind, focusing specifically on the use of two or more moderators to maximize the predictability of virtually all the Ss in the cross-validation samples.

REFERENCES

- BERDIE, R., & LAYTON, W. *Manual, Minnesota Counseling Inventory*. New York: Psychological Corporation, 1957.
- BUREAU OF RESEARCH, WISCONSIN STATE DEPARTMENT OF HEALTH AND SOCIAL SERVICES. *Juveniles admitted to Division of Correction's juvenile institutions in 1966*. (Statistical Bulletin C-51) Madison: Author, June 1967.
- BUREAU OF RESEARCH, WISCONSIN STATE DEPARTMENT OF HEALTH AND SOCIAL SERVICES. *Juveniles released from Division of Correction's juvenile institutions in 1966*. (Statistical Bulletin C-52) Madison: Author, September 1967.
- COWDEN, J. Predicting institutional adjustment and recidivism in delinquent boys. *Journal of Criminal Law, Criminology and Police Science*, 1966, 57, 39-44.
- COWDEN, J., & PACHT, A. Predicting institutional and postrelease adjustment of delinquent boys. *Journal of Consulting Psychology*, 1967, 31, 377-381.
- GHEISELLI, E. Differentiation of individuals in terms of their predictability. *Journal of Applied Psychology*, 1956, 40, 374-377.
- GHEISELLI, E. The prediction of predictability. *Educational and Psychological Measurement*, 1960, 20, 3-8.
- GHEISELLI, E. Moderating effects and differential reliability and validity. *Journal of Applied Psychology*, 1963, 47, 81-86.

(Received September 23, 1968)

SEX DIFFERENCES AND VARIATIONS IN BODY ATTITUDES

RICHARD M. KURTZ¹

Washington University

In the present study a number of hypotheses were tested about the relationship between body attitude and sex membership. The Ss were 89 male and 80 female undergraduate students. Three separate body attitudes corresponding to Osgood's three attitudinal dimensions were measured by the Semantic Differential on 30 different body concepts. Results indicated that women not only like their bodies better than men but also have a more clearly differentiated notion of what they like and dislike about their bodies. Men on the other hand judge their bodies as more potent and more active than women.

The purpose of this study was to investigate the relationship between sex differences and variations in global body attitudes. A body attitude is defined as an individual's general, overall global attitude or feeling about the outward form and appearance of his body. Three separate body attitudes will be studied which correspond to the three major factors Osgood, Suci, and Tannenbaum (1957) have isolated (Evaluation, Potency, and Activity).

Of all the differences that exist between human beings, sex membership is presumably one of the most important. Sex and age are equally universal as reference points for establishing status and assigning social roles (Lundberg, Schrag, & Larsen, 1954). Most societies prescribe different attitudes and activities to and develop different expectancies about men and women. The techniques used to effect and maintain such differences also vary from culture to culture (Mead, 1935, 1949). In the ascription of occupation, education, political activity, and moral codes, sex difference is an important factor in most societies.

In light of the presumed general importance of sex membership, it is indeed surprising that there are few experimental studies which deal directly with sex differences and variations in body attitude. Except for the various psychoanalytic speculations about the importance of the anatomical differences between the sexes (Freud, 1938, 1959b, 1959c), there is only one study directly related to possible differences in body concepts as a function of sex

membership. Jourard and Remy (1957) attempted to investigate whether men and women differed in their degree of differentiation in global body attitude. They tested the differentiation hypothesis² by comparing the between-person variances of the males and females in terms of their total body attitude scores. No definite conclusions can be drawn from their study since comparison of between-person variance is not an adequate test of the differentiation hypothesis. In other words, insofar as between-person variance reflects only differences between the groups in overall attitude, it reflects only variance due to individual differences in total body attitude. The correct test would have been to compare the within-person variances of the males and females.

Several studies are distantly related to the question of sex differences in body attitudes. For example, some researchers (Secord, 1953; Weinberg, 1960) have found that women gave more homonym associations to body concepts than did men. In a similar study using the TAT, Van Lennep (1957) reported that females exceed males in the number of body-related references found in their stories. In a study with male and female adolescents, Erikson (1951) found differences in the type of scenes which male and female adolescents

² Differentiation refers to intraindividual variation in attitudes toward specific body aspects. The concept of differentiation is identical with the semantic distance notion of Osgood et al. (1957). The more articulated a notion of what one likes or dislikes about his body, the clearer the evaluative distinctions a person can be expected to draw. The greater the differentiation, the greater will be the differences between all body concepts.

¹ Requests for reprints should be sent to Richard M. Kurtz, Department of Psychology, Washington University, St. Louis, Missouri 63130.

constructed with play material. Females generally made structures which were open and easy to enter, whereas males fashioned more upright "phallic" type structures. Erikson speculated that these differences reflected contrasting modes of experiencing their bodies which were related to their over all sexual differentiation. Using a task which called for the construction of a series of human forms by means of schematic body parts of different sizes, Katcher and Levin (1955) found that girls were able to match body parts to form representing self, mother, and father more accurately than could a group of boys of the same age. They interpreted their data to mean that girls may have a more realistic appreciation of the smallness of their own bodies in relation to adults than is the case with boys.

While these studies provide some fragmentary and tangential support for the existence of sex differences in body attitude, most of them serve even more to highlight the need for a direct "common sense" approach to the investigation of body attitude. It seems to this author that, in American society, differential attitudes toward the body are readily apparent. General and specific aspects of the female body, for instance, are held in high esteem. One need only glance through a few of countless magazines, walk past movie marquees, beauty shops, or cosmetic counters to verify the importance placed on the female body. It is also a commonplace observation that women expend more effort, time, and money on personal grooming than do men.

One hypothesis suggested by these considerations is that women should have a more clearly differentiated or articulated notion of what aspects they like and what aspects they dislike about their bodies than is generally the case with men. Females can, therefore, be expected to draw finer evaluative distinctions about the various aspects of their bodies and might be less inclined to respond to their bodily appearance in a global manner than men. If this is the case then one would predict that all possible differences between all body concepts produced by females on the Evaluative factor of the Semantic Differential will be larger than all possible differences between the same concepts produced by males on the Evaluative factor. Stated in an equiva-

lent but more technical way, it is predicted that the within-person variance of the females' Evaluative factor scores will be significantly larger than the within-person variance of the males' Evaluative scores.

It is also predicted that women will evaluate their bodies more positively than will men. That sex differences are related to the attitudes and expectancies held about the body is also indicated by another common sense observation. Muscular strength (Jones, 1947) is an admired quality in men, but not in women. Aggressive behavior is less acceptable in females than in males (Lansky, Grandall, Kagan, and Baker, 1961). In interpersonal relationships the man is expected to be strong and dominant, the woman weak and submissive. Freud (1933, 1959a,) has seen the failure to accept these role prescriptions as involved in the genesis of neurosis. From these observations, the following hypothesis is offered: Males will tend to judge their bodies as more potent or stronger than will females, that is, males will rate their bodies higher on the Potency factor than will females.

Sex differences can also be expected on the Activity dimension of the Semantic Differential. By the time girls reach puberty, most of them accept certain restrictions on their physical activities (Muuss, 1962). They are exposed to strong social pressures to relinquish, among other things, tomboy behavior. Most girls, at least those from the middle class, come to recognize that it is not ladylike to be as rough and active as their brothers. A little later on, girls also learn that in social situations involving the opposite sex, it is unacceptable to be "forward," that is, to take the initiative, be it sexual or otherwise. The female, in short, is generally assigned a more passive role in many situations in our society than is the case with males.

Women's acceptance of a more passive role is seen by many psychiatrists (Deutsch, 1930; Fenichel, 1955) as a precondition of sexual adjustment and marital happiness. On the other hand, the reluctance or inability of a man to be assertive and accept the active role has been viewed as pathogenic in the formation of sexual impotence and homosexuality (Ferenczi, 1950a, 1950b). For these reasons, a difference is expected between males and

TABLE 1

THIRTY BODY CONCEPTS JUDGED BY SUBJECTS ON THE
SEMANTIC DIFFERENTIAL IN ORDER OF
PRESENTATION

Color of my hair
Size of my hands
Shape of my fingers
Size of my ears
Shape of my ankles
Shape of my head
My profile
Width of my shoulders
Size of my bust (females)
Size of my chest (males)
Size of my hips
Shape of my legs
Size of my feet
My posture
My weight
Texture of my hair
My facial complexion
Distribution of hair on my body
Size of my waist
Shape of my ears
Length of my neck
My body build
My height
Size of my arms
Appearance of my facial features
Texture of my skin
Appearance of my teeth
Shape of my knees
My gait
Back view of my head
Color of my skin

females in the way they rate their bodies on the Activity factor. College men will judge their bodies as more active than will college women.

METHOD

The Ss were 89 males and 80 female undergraduate Caucasians, predominantly middle-class students between 18 and 23 years old.

Three separate body attitudes corresponding to Osgood's three attitudinal dimensions (Evaluation, Potency, and Activity) were measured by the Semantic Differential on 30 different body concepts, such as "size of my arms," "color of my hair." The format for presenting this technique to Ss was identical to that of Osgood. The body concepts are listed in Table 1.

The task of S was to rate the body concept in question along nine bipolar 7-point adjective scales^a

^a The nine bipolar 7-point adjective scales used in this study were: Evaluative dimension, good-bad, awkward-graceful, beautiful-ugly; Potency dimension, weak-strong, hard-soft, thin-thick; Activity dimen-

which tap the three attitude dimensions. The sum of the judgments on the three adjective scales for each dimension of the Semantic Differential were summed separately for each concept to obtain an item score. Then the 30 item scores were summed for each person and the total sum was treated as an index of how a person felt about the overall appearance of his body.

A generalizability study (Gleser, Cronbach, & Rajaratnam, 1965) was conducted on a pilot basis to see how well one could generalize from attitudes toward specific aspects of the bodily appearance to a hypothesized universe of observations, the global body attitude. The "wanted" variance due to differences among Ss over a universe of body attitudes and adjectival scales far exceeded the "unwanted" variance in the total scale score attributable to idiosyncratic responses of Ss to specific concepts on specific scales. The obtained generalizability coefficients for all three dimensions ranged from .86 to .95.

RESULTS

The hypothesis that women would have a more clearly differentiated notion of what they like or dislike about their bodies than would men was confirmed. An *F* ratio computed between the within-person mean square for women ($MS = 11.851$) and the within-person mean square for men ($MS = 8.307$) yielded an *F* value of 1.425 which was significant beyond the .05 level.

In addition to having a more differentiated body concept, females also like their bodies better than do men. The test of significance for differences between the mean for females ($\bar{X} = 452.01$) and the mean for males ($\bar{X} = 430.03$) yielded a *z* of 2.27 significant beyond the .02 level.

The hypothesis that the mean Potency factor score for men ($\bar{X} = 414.72$) would be significantly larger than the mean Potency factor score for women ($\bar{X} = 351.11$) was confirmed. A test of significance of differences between these means yields a *z* = 7.47 significant beyond the .001 level.

The hypothesis which predicted that the mean Activity factor score for men ($\bar{X} = 419.69$) would be significantly larger than the mean Activity factor score for women ($\bar{X} = 403.35$) was confirmed. The test for significance yields a *z* = 2.20 which is significant

sion, active-passive, cold-hot, fast-slow. The poles of the scales and the position of the dimensions were systematically counterbalanced in order of presentation.

beyond the .02 level. Pearson product-moment correlation coefficients computed separately for both sexes between the three attitude dimensions ranged from .01 to .09 and were not significantly different from zero, thus indicating that all three factors were orthogonal.

DISCUSSION

In the present study the hypothesis was confirmed that college women have a more clearly differentiated body concept than men. This finding suggests that greater awareness and concern over bodily appearance may be more acceptable in the female than in the male. This differentiation may stem from the fact that it is part of the females' role prescription to focus attention on the details of their bodies, whereas men are expected to be more subdued in the interest they take in the appearance of their bodies.

The finding that women have a more positive evaluative attitude toward their bodies than men is consistent with the high esteem placed on the physical form of the female body in American society. It is interesting to contrast these findings with the psychoanalytic contention (Kubie, 1937; Reik, 1957) that women have a lower evaluation of their bodily appearance than men because of "penis envy." Sex differences in body concept may, of course, be related to more than obvious anatomical differences. It is possible that particular cultural definitions of masculinity and femininity may be related to specific sex differences found in the three factors of the Semantic Differential. Thus, in a culture which emphasized the need for the female to be physically strong and assertive and the male weak and passive, there might be a reversal of the present findings. One might even speculate that greater body differentiation and high evaluation in the female is related to her lack of actual social or political power in American society. If this is true one might expect differences in body attitude between women who are occupationally successful in a "man's world," and those women who are content to accept a more traditional female role. The career women, therefore, might have a body attitude profile more closely resembling the typical male, that is, she might be lower

on Evaluation and higher on Potency and Activity, than would be the more traditionally oriented woman.

On the other hand, it may be that the actual physical differences in the form and appearance of the male and female bodies may be the nucleus for the differences in body attitude (Freud, 1933, 1959b). In this case one would expect similar sex differences in body attitude regardless of cultural variations in role prescriptions. These speculations about the underlying factors which produce sex differences in body attitude pose questions for future research.

Sex differences in body attitude may also provide useful leads for studying several theoretical problems about sex roles within each sex. Sex differences in the multi-dimensional body attitude measure may, for example, reflect general feelings of masculinity or femininity. Perhaps body attitude is related to sexual adjustment. Do women who judge their bodies as very potent and very active have greater sex role conflict than women who see their bodies as weak and passive? Similar questions can be raised for men. Are men who judge their bodies as weak and passive more inclined to homosexual or latent homosexual adjustments?

Thus, it seems possible that measures of body attitude can provide an additional tool to probe the masculinity-femininity dimension. This area is one that might prove very fruitful for future investigation since what counts as feminine or masculine in American culture so often seems associated with what may loosely be called a "good figure." Association of the masculinity-femininity dimension with body attitude will probably be most pronounced in the adolescent and young adult age range. One might also expect social class differences here, with the lower classes relying more heavily upon a "good physique" to define and support the adequacy of their sex role identity.

REFERENCES

- DEUTSCH, H. The significance of masochism in the mental life of women. *International Journal of Psychoanalysis*, 1930, 11, 48-58.
- ERIKSON, E. Sex differences in the play configurations of pre-adolescents. *American Journal of Orthopsychiatry*, 1951, 21, 667-689.

- FERENCZI, O. Further light upon the pre-oedipal phase in girls. In, *Collected papers of Otto Ferenczi: First series*. New York: Norton, 1955.
- FERENCZI, S. The analytic interpretation and treatment of psychosexual impotence. In S. Ferenczi, *Sex and psychoanalysis*. New York: Basic Books, 1950. (a)
- FERENCZI, S. The nosology of male homosexuality. In S. Ferenczi, *Sex and psychoanalysis*. New York: Basic Books, 1950. (b)
- FREUD, S. Three contributions to the theory of sex. In, *The basic writings of Sigmund Freud*. New York: Random House, 1938.
- FREUD, S. The psychogenesis of a case of homosexuality in a woman. In, *Collected papers*. Vol. 2. New York: Basic Books, 1959. (a)
- FREUD, S. Some psychological consequences of the anatomical distinction between the sexes. In, *Collected papers*. Vol. 5. New York: Basic Books, 1959. (b)
- FREUD, S. Female sexuality. In, *Collected papers*. Vol. 5. New York: Basic Books, 1959. (c)
- FREUD, S. The psychology of women. In, *New introductory lectures on psychoanalysis*. New York: Norton, 1933.
- GLESER, G., CRONBACH, L., & RAJARATNAM, N. Generalizability of scores influenced by multiple sources of variance. *Psychometrika*, 1965, 30, 395-418.
- JONES, H. Strength and physique. *American Journal of Physical Anthropology*, 1947, 5, 29-39.
- JOURARD, S., & REMY, R. Individual variance score: An index of the degree of differentiation of the self and the body image. *Journal of Clinical Psychology*, 1957, 13, 62-63.
- KATCHER, A., & LEVIN, M. Children's conceptions of body size. *Child Development*, 1955, 26, 103-110.
- KUBIE, L. The fantasy of dirt. *Psychoanalytic Quarterly*, 1937, 6, 388-424.
- LANSKY, L., GRANDALL, V., KAGAN, J., & BAKER, C. Sex differences in aggression and its correlates in middle-class adolescents. *Child Development*, 1961, 32, 45-58.
- LUNDBERG, G., SCHRAG, C., & LARSEN, O. *Sociology*. New York: Harper, 1954.
- MEAD, M. *Sex and temperament in three primitive societies*. New York: Morrow, 1935.
- MEAD, M. *Male and female*. New York: Morrow, 1949.
- MUUSS, R. *Theories of adolescence*. New York: Random House, 1962.
- OSGOOD, C., SUCI, G., & TANNENBAUM, P. *The measurement of meaning*. Urbana: University of Illinois Press, 1957.
- REIK, T. The emotional differences of the sexes. In T. Reik (Ed.), *Of love and lust*. New York: Grove Press, 1957.
- SECORD, P. Objectification of word-association procedures by the use of homonyms: A measure of body cathexis. *Journal of Personality*, 1953, 21, 479-495.
- VAN LENNER, D. Projection and personality. In H. David & E. Von Bracken (Eds.), *Perspective in personality theory*. New York: Basic Books, 1957.
- WEINBERG, J. A further investigation of body cathexis and the self. *Journal of Consulting Psychology*, 1960, 24, 277.

(Received October 4, 1968)

BRIEF REPORTS

EFFECT OF INTELLIGENCE, ANXIETY, AND DIAGNOSIS ON ARITHMETIC AND DIGIT SPAN PERFORMANCE ON THE WAIS¹

CHARLOTTE J. CALLENS² AND MALCOLM L. MELTZER

George Washington University

The purpose of this study was to investigate the effects of intelligence, anxiety, and diagnosis on the relationship between the Arithmetic (A) and Digit Span (DS) subtests of the WAIS. Of special interest was the hypothesis of Rapaport, Gill, and Schafer (1945) that schizophrenics show the "out of pattern" relationship of DS higher than A as compared to the more typical A higher than DS found in the anxious normal or neurotic.

Eighty male Ss were selected from a larger pool of patients who had been hospitalized at either a large city hospital or a naval hospital. Forty carried discharge diagnoses of schizophrenia, 40 were diagnosed as neurotic. WAIS scores with A and DS subtracted were used to divide the sample into high IQ (105 or above) and low IQ (104 or below). The Welsh A scale of the MMPI was employed as a measure of anxiety (*T* score 55 or above) and low anxiety (*T* score 54 or below). This allowed for a $2 \times 2 \times 2$ factorial design composed of eight cells and with 10 Ss per cell. Analysis of variance was used to evaluate effects of intelligence, anxiety, and diagnosis on A, DS, and A minus DS.

Intelligence was found to be related to A performance ($F = 24.8$, $df = 1/72$, $p < .001$) but

not to DS ($F = 1.79$). Higher A than DS score was also related to intelligence ($F = 9.04$, $df = 1/72$, $p < .01$).

No direct relationship was found between anxiety and either A or DS scores. However, the difference between A and DS was related to anxiety, with high anxiety being associated with better A performance than DS performance ($F = 6.02$, $df = 1/72$, $p < .02$).

Diagnosis did not have a direct effect on A, DS, or A minus DS.

The Intelligence \times Anxiety \times Diagnosis interaction approached significance ($F = 3.61$, $df = 1/72$, $p < .10$). None of the low-anxiety, low-IQ schizophrenics showed A superior to DS, whereas this was the typical pattern for the high-anxiety, high-IQ neurotics. The *t* test was significant at the .001 level ($t = 4.73$, $df = 18$). However, the A to DS relationship is clearly a function of intelligence and anxiety, and the DS above A pattern is to be expected in either diagnostic group when S is of low-intelligence or low-anxiety level.

The findings of this study suggest that to interpret the relationship between A and DS one must consider the intelligence level of S, and that the relationship between A and DS is probably a better indication of anxiety than is the level of DS alone. DS well below A may be a function of anxiety or intelligence, and the effect would be enhanced when S is both bright and anxious. Rapaport, Gill, and Schafer (1945) are probably right in assuming that the chronic or low-anxiety, dull schizophrenic will show the out-of-pattern relationship of DS above A, but so will nonschizophrenics who are dull or not anxious.

REFERENCE

- RAPAPORT, D., GILL, M. M., & SCHAFER, R. *Diagnostic psychological testing*. Vol. 1. New York: Year Book Publishers, 1945.

(Received February 13, 1969)

¹This paper is based on a thesis submitted by the first author in partial fulfillment of the requirements for the MA degree at the George Washington University. The study was carried out under the direction of Malcolm L. Meltzer and Bernard I. Levy.

An extended report of this study may be obtained without charge from Malcolm L. Meltzer, Department of Psychology, George Washington University, Washington, D. C. 20006, or for a fee from the American Society of Information Science. Order NAPS Document 00579 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 909 Third Ave., New York, New York 10022; remitting \$1.00 for microfiche or \$3.00 for photocopies. Requests for reprints should be sent to Malcolm L. Meltzer at the above address.

²Now at Prince Georges Community College.

FIELD INDEPENDENCE AND EDUCATIONAL-VOCATIONAL INTERESTS¹

JACK ARBUTHNOT AND LEOPOLD GRUENFELD

Cornell University

The purpose of this study was to determine whether field dependence-independence (FD-FI) is predictive of educational-vocational interests. It is hypothesized that (a) increasing FI will be associated with increasing interest in analytical and abstract areas, such as physics, and with decreasing interest in less analytical and abstract areas, such as the humanities; (b) since relative masculinity has been shown to modify relationships involving FI (see Gruenfeld & Arbuthnot, 1969) for male Ss, the relationships predicted in a will be of greater magnitude for more masculine Ss; (c) the partialing out of effects due to verbal and mathematical intelligence will not affect the magnitude of the relationships predicted in a.

The sample consisted of 55 male freshmen engineering students. Level of FI was assessed by both the portable rod-and-frame test (RFT) and the group form Embedded Figures test (GEFT; Form CFV). Masculinity was assessed by the Guilford-Zimmerman M-F scale (GZMF). Verbal and mathematical intelligence were assessed by the Scholastic Aptitude Test (VSAT and MSAT). Interests were measured by the Business (BUS), Humanities (HUM), Social (SOC), Physical (PHY), and Biological (BIO) Sciences scales of Fricke's (1965) Opinion, Attitude, and Interest Survey. Analytical interests are considered to be PHY (most pure) and BIO; nonanalytical are HUM (most pure), BUS, and SOC.

The results of Pearson product-moment correlational analyses were as follows: The RFT

and GEFT were independent ($r = -.03$) and lacked convergent validity; this replicates earlier findings of Gruenfeld and Arbuthnot. The GZMF was independent of RFT but predicted HUM, PHY, and BIO ($r = -.42, .32, .32, p < .05$). The interest scales intercorrelated negatively or not at all. Regarding the hypotheses, RFT (but not GEFT) predicted PHY, BIO, HUM, and SOC ($r = -.33, -.28, .34, .32, p < .05$) in the expected directions. With the effects due to VSAT partialled out, the relationships of RFT to PHY and HUM (the purest scales) remained significant, while BIO and SOC did not; MSAT was independent of RFT and all interest scales. When the 38 Ss who were given the GZMF were split into high ($N = 18$) and low ($N = 20$) masculine groups, results opposite to those hypothesized were observed. For the low-masculine Ss, RFT (but not GEFT) predicted PHY, HUM, and BUS ($r = -.56, .50, .45, p < .05$). With VSAT partialled out, no changes in significance were observed; MSAT was independent of RFT and all interest scales.

The results strongly support the hypothesis that FD-FI (as measured by RFT) is predictive of educational-vocational interests. Further, the notion that FD-FI merely reflects intelligence is largely dispelled for at least one portion of the population. Finally, the GZMF does not simply reflect predictability; it may be reflecting either a rigidity-flexibility or an instrumental-expressive dimension, particularly for FI Ss, but which is independent of FI. Further research on both the prediction of interests with a larger, more representative sample and the nature of the GZMF moderating effect is underway.

REFERENCES

- FRICKE, B. *Opinion, Attitude and Interest Survey handbook*. Ann Arbor: OAIS Testing Program, 1965.
- GRUENFELD, L., & ARBUTHNOT, J. Field independence as a conceptual framework for the prediction of variability in the rating of others. *Perceptual and Motor Skills*, 1969, 28, 31-44.

(Received June 4, 1969)

¹ An extended report of this study may be obtained without charge from Jack Arbuthnot, New York State School of Industrial and Labor Relations, Cornell University, Ithaca, New York 14850, or for a fee from the American Society for Information Science. Order NAPS Document 00578 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 909 Third Avenue, New York, New York 10022; remitting \$1.00 for microfiche or \$3.00 for photocopies. Requests for reprints should be sent to Jack Arbuthnot at the above address.

IMPACT OF AUDIO-VISUAL FEEDBACK ON GROUP PSYCHOTHERAPY¹

BURTON N. DANET

Hamm Memorial Psychiatric Clinic, St. Paul, Minnesota

In this study, also described elsewhere (Danet, 1968, 1969), a beginning attempt to assess the impact of videotape playback on group therapy under relatively controlled conditions was made. Two university student groups from the Mental Hygiene Clinic, University of Minnesota Health Service, were videotaped under identical conditions. The experimental group (E) received playback while the control group (C) did not. C contained four men and three women, E the reverse. The same co-therapists conducted both groups.

The groups met weekly for ten 50-minute sessions. For E, 10-minute segments were extracted from the previous meeting for playback at the beginning of the following session.

A periodic questionnaire was administered. Sociometric items dealing with interpersonal attractions were aimed at assessing interaction processes and cohesiveness. Evaluative ratings and a mutual ranking by members on contribution to the group's progress were also included. The ranking yielded an "index of consensus on rankings in small groups" which made it possible to compare their achievement of concordance. After the last session, students rated other members regarding the extent to which they had benefited from therapy. Two other procedures employed were an analysis of the seating arrangements across sessions and a frequency count of group members' participation.

¹ This report is based on a PhD dissertation, University of Minnesota, 1967. The author wishes to thank Robert D. Wirt, dissertation advisor, for his encouragement. The research was supported by a grant from the General Research Fund, University of Minnesota Medical School.

An extended report of this study may be obtained without charge from Burton N. Danet, who is now at the William Alanson White Institute, 20 West 74th Street, New York, New York 10023, or for a fee from the National Auxiliary Publications Service. Order NAPS Document 00577 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 909 Third Avenue, New York, New York 10022; remitting \$1.00 for microfiche or \$3.00 for photocopies. Requests for reprints should be sent to Burton N. Danet at the above address.

The results were as follows: In E the chairs were occupied by a significantly larger number of students across sessions than in C. This finding of a greater amount of mobility and shifting position from session to session was consistent with Geertsma and Reivich's (1965) assumption that any self-observational experience is potentially anxiety producing.

A tally of the number of first choices each member received on the sociometric questions showed that C "endorsed" one or two members and failed to endorse one or two others, while in E, first choices tended to fluctuate. Results of the evaluative and improvement ratings showed that E tended to have a more negative view of itself.

The frequency of participation data revealed C as a well-defined group with a stable pattern of interaction. No pattern was apparent in E which was more variable. Only on the task-type dimension of the mutual ranking on contribution to the group's progress did E achieve a higher degree of concordance (unity).

The results pointed to the a posteriori hypothesis that the videotape playback as provided during this study had a disrupting influence. As compared to C, E was hampered in establishing itself as a productive, cohesive unit. This finding supported Stoller's (1967) warning that if videotape feedback is not handled adequately by a skillful therapist, it can be disruptive of a psychotherapy group's progress.

REFERENCES

- DANET, B. N. Self-confrontation in psychotherapy reviewed. *American Journal of Psychotherapy*, 1968, 22, 245-257.
- DANET, B. N. Video tape playback as a therapeutic device in group psychotherapy. *International Journal of Group Psychotherapy*, 1969, in press.
- GEERTSMA, R. H., & REIVICH, R. Repetitive self-observation by videotape playback. *Journal of Nervous and Mental Disease*, 1965, 141, 29-41.
- STOLLER, F. H. Focused feedback with videotape: Extending the group's function. In G. M. Gazda (Ed.), *Basic innovations in group therapy and counseling*. Springfield, Ill.: Charles C Thomas, 1967.

(Received January 15, 1969)

CONSULTING PSYCHOLOGIST

Mature Ph.D. Psychologist with strong clinical skills. Should be experienced counseling with mature adults. Supervisory experience desirable. Prefer knowledge of guidance and industrial counseling.

To counsel executives and management personnel in areas of personal development and career planning. Handle assignments in assessment, training and organization.

Starting salary \$15,000 to \$25,000 based upon experience and qualifications. Attractive benefits, growth opportunity. Little or no travel.

Opportunities in New York, Chicago, Dallas, Los Angeles, San Francisco and other major cities.

For consideration phone or submit resume in confidence:

C. T. Deeble, Ph.D.

American Management Psychologists

**200 South Michigan Avenue, Chicago, Illinois 60604
312-427-9335**

Journals of the AMERICAN PSYCHOLOGICAL ASSOCIATION

Subscription Rates for Calendar Year 1970

AMERICAN PSYCHOLOGIST. Editor: Charles S. Gersoni, American Psychological Association. Official papers of the Association and articles on Psychology. Monthly. The 1970 volume is Vol. 25. Subscription: \$10.00 (foreign \$10.50). Single copy \$1.00.

CONTEMPORARY PSYCHOLOGY. Editor: Gardner Lindzey, University of Texas. Critical reviews of books, films, and research material in the field of psychology. Monthly. The 1970 volume is Vol. 15. Subscription: \$10.00 (foreign \$10.50). Single copy \$1.00.

JOURNAL OF ABNORMAL PSYCHOLOGY. Editor: Howard F. Hunt, Columbia University. Theoretical and research conditions in behavior pathology. Bimonthly; first issue appears in February. Beginning 1970. Two volumes per year, three issues per volume. The 1970 volumes are Vols. 75 and 76. Subscription: \$20.00 for 2 vols. (foreign \$20.50). Single copy \$4.00.

JOURNAL OF APPLIED PSYCHOLOGY. Editor: Kenneth E. Clark, University of Rochester. Applications of psychology in business and industry. Bimonthly; first issue appears in February. The 1970 volume is Vol. 54. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF COMPARATIVE AND PHYSIOLOGICAL PSYCHOLOGY. Editor: Elliot Stellar, University of Pennsylvania. Original contributions in the field of comparative and physiological psychology. Monthly. Beginning in 1970: four volumes per year, three issues per volume. The 1970 volumes are Vols. 70, 71, 72 and 73. Subscription: \$40.00 or 4 vols. (\$40.50 foreign). Single copy \$4.00.

JOURNAL OF COUNSELING PSYCHOLOGY. Editor: Francis P. Robinson, Ohio State University. A primary publication medium for research in counseling theory and practice. Bimonthly; first issue appears in January. The 1970 volume is Vol. 17. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF CONSULTING AND CLINICAL PSYCHOLOGY. Editor: Jules D. Holzberg, Wesleyan University. Research in clinical psychology: psychological diagnosis, psycho-therapy, personality, psychopathology. Bimonthly; first issue appears in February. Beginning in 1970: Two volumes per year, three issues per volume. The 1970 volumes are 34 and 35. Subscription: \$20.00 for 2 vols. (foreign \$20.50). Single copy \$4.00.

JOURNAL OF EDUCATIONAL PSYCHOLOGY. Editor: Wayne H. Holtzman, University of Texas. Studies of learning and teaching, measurement of psychological development, psychology of school subjects, methods of instruction, school adjustment. Bimonthly; first issue appears in February. The 1970 volume is Vol. 61. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY. Editor: David A. Grant, University of Wisconsin. Original contributions of an experimental character. Monthly. Four volumes per year, three issues in each volume. The 1970 volumes are Vols. 83, 84, 85 and 86. Subscription: \$10.00 for 4 vols. (foreign \$40.50). Single copy \$4.00.

JOURNAL OF PERSONALITY AND SOCIAL PSYCHOLOGY. Editor: William J. McGuire, University of California. Theoretical and research papers on personality dynamics, group process, and the psychological aspects of structure. Monthly. Three volumes per year, four issues in each volume. The 1970 volumes are Vols. 14, 15 and 16. Subscription: \$30.00 for 3 vols. (foreign \$30.50). Single copy \$3.00.

PSYCHOLOGICAL ABSTRACTS. Executive Editor: Philip J. Siegmund, American Psychological Association. Noncritical abstracts of the world's literature in psychology and related subjects. Monthly and an Annual Cumulative Index. Approximately 20,000 abstracts published annually. The 1970 volume is Vol. 44. Subscription: \$130.00 (\$132.00 foreign).

PSYCHOLOGICAL BULLETIN. Editor: James Deese, Johns Hopkins University. Evaluative reviews of research literature, discussions of research methodology in psychology. Monthly. Two volumes per year, six issues in each volume. The 1970 volumes are Vols. 73 and 74. Subscription: \$20.00 for 2 vols. (foreign \$20.50). Single copy \$2.00.

PSYCHOLOGICAL REVIEW. Editor: Charles N. Cofer, Pennsylvania State University. Original contributions of a theoretical nature. Bimonthly; first issue appears in January. The 1970 volume is Vol. 77. Subscription: \$10.00 (foreign \$10.50). Single copy \$2.00.

DEVELOPMENTAL PSYCHOLOGY. Editor: Boyd R. McCandless, Emory University. Studies on growth and development and their associated variables. Bimonthly; first issue appears in January. Beginning in 1970: Two volumes per year, three issues per volume. The 1970 volumes are Vols. 2 and 3. Subscription: \$20.00 for 2 vols. (\$20.50 foreign). Single copy \$4.00.

Discounts: AGENTS and BOOK DEALERS receive a 10% discount on all orders. There is no club arrangement. There are no quantity, cash, library, or other special discounts. *Payment in U. S. funds.*

★ Subscriptions are available on a January through December basis only ★

AMERICAN PSYCHOLOGICAL ASSOCIATION
Subscription Department
1200 Seventeenth Street, N.W.
Washington, D. C. 20036

JOURNAL OF CONSULTING and CLINICAL PSYCHOLOGY

December 1969

Vol. 33, No. 6

Contents

ARTICLES

Manipulation of Indications of Disturbed Thinking in Normal Subjects: <i>Leonard S. Levitz and Leonard P. Ullmann</i>	633
Schizophrenia as a "Put-On": <i>Gerald Rosenbaum</i>	642
Schizomimetic Conditions and Schizophrenia: <i>Loren J. Chapman</i>	646
An Experimental Procedure for the Modification of Psychotic Behavior: <i>Ivan Toby Rutner and Charles Bugle</i>	651
Perception of Disability by the Nondisabled: <i>A. P. MacDonald, Jr., and Janet Hall</i> ..	654
Frequency of Dream Recall Estimated by Three Methods and Related to Defense Preference and Anxiety: <i>David Benjamin Cohen</i>	661
Intelligence Defined as Effectiveness of Approaches: <i>Constance T. Fischer</i>	668
✓Operant-Interpersonal Treatment for Marital Discord: <i>Richard B. Stuart</i>	675
✓Personal Adjustment as a Determinant of Aggressive Behavior toward the Mentally Ill: <i>Kenneth Ring and Amerigo Farina</i>	683
Effect of Information, Subject Age, and Instructional Set on Predictive Judgments: <i>James L. Knestrick and Leon Gorlow</i>	691
✓Rorschach Prediction of Success in Clinical Training: A Second Look: <i>Rae Carlson</i> ...	699
✓Verbal and Overt-Behavioral Assessment of a Specific Fear: <i>Anthony F. Fazio</i>	705
Relations between Personality and Smoking Behavior in Preadult Subjects: <i>Gene M. Smith</i>	710
Effects of Cues and Examiner Influence on Two Wechsler Subtests: <i>Jerome M. Sattler</i>	716
Personality Correlates of Accurate Empathy in a College Companion Program: <i>George J. Vesprani</i>	722
Identification of Clinically Relevant Dimensions of Children's Behavior: <i>Jacob O. Sines, Jerome D. Pauker, Lloyd K. Sines, and David R. Owen</i>	728
Relationship of Work Therapy to Psychiatric Length of Stay and Readmission: <i>Margaret S. Barbee, K. L. Berry, and Leo A. Micek</i>	735
Some Personality and Aptitudinal Correlates of the "A-B" Therapist Scale: <i>James E. Dublin, Charles F. Elton, and Juris I. Berzins</i>	739
Distinctive Personality Attributes of Child-Abusing Mothers: <i>Barry Melnick and John R. Hurley</i>	746
SPI Discriminators among Four Psychological Health-Sickness Levels: <i>Don Spiegel</i> ..	750
Prospective Span as a Cognitive Ability: <i>Louis S. Dickstein</i>	757
Should WISC Subjects Explain Picture Arrangement Stories?: <i>James M. Herrell and Jeffrey H. Golland</i>	761
Relationship of Manifest Anxiety to Repression-Sensitization on the MMPI: <i>Patrick F. Sullivan and Lynn K. Roberts</i>	763
LIST OF MANUSCRIPTS ACCEPTED.....	738, 760

(Volume title page and contents appear at the end of this issue.)

Brief Reports

The *Journal of Consulting and Clinical Psychology* will accept Brief Reports of research studies in clinical psychology for early publication without expense to the author. The procedure is intended to permit the publication of soundly designed studies of specialized interest or limited importance which cannot now be accepted because of lack of space. Several pages in each issue will be devoted to Brief Reports, published in the order of their receipt without respect to the dates of receipt of the regular articles. Most Brief Reports appear in the first or second issue to go to press following their final acceptance.

An author who wishes to submit a Brief Report:

1. Sends the Brief Report, limited to one printed page and prepared according to the specifications given below.

2. Also sends to the Editor a full report of the research study, in sufficient detail to give a clear account of its background, procedure, results, and conclusions, which will be filed with the National Auxiliary Publications Service (NAPS) to insure indefinite availability.

3. Prepares at least 100 mimeographed copies of the full report, which the author will send without charge to all who request it as long as the supply lasts.

4. Agrees not to submit the full report to another journal of general circulation.

Specifications

Brief Report. The Brief Report should give a clear, condensed summary of the procedure of the study and as full an account of the results as space permits.

To insure that the Brief Report will be no longer than one printed page, its typescript, in-

cluding all matter except the title and the author's lines, must not exceed 85 lines averaging 42 characters and spaces in length. Set the typewriter margins for short lines of 42 characters, which are 3.5 inches long in elite typing, and 4.2 inches long in pica.

The manuscript of the Brief Report must be *double spaced* throughout. Except for its short lines, it follows the standard style of the 1967 revision of the *APA Publication Manual*. Headings, tables, and references are avoided or, if essential, must be counted in the 85 lines. Each Brief Report must be accompanied by a footnote in the style below, which is typed on a separate sheet and *not* counted in the 85-line quota:

¹ An extended report of this study may be obtained without charge from John Doe (giving the author's full name and address) or for a fee from the American Society for Information Science. Order NAPS Document — from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 909 Third Avenue, New York, New York 10022; remitting \$— for microfiche or \$— for photocopies.

Extended report. Because the extended report is intended for photoduplication, and is not copy to be sent to a printer, its style should differ in several ways from that of other manuscripts: (a) The extended report should be typed with single spacing for economy in duplication. (b) Tables and figures should be placed adjacent to the text which refers to them. A caption should be typed below each figure. (c) Footnotes should be typed at the bottom of the page on which reference is made to them. In other respects, the full report is prepared in the style specified by the *Publication Manual*.

Style of Manuscripts. Manuscripts must be prepared in the style described in the *Publication Manual of the American Psychological Association (1967 Revision)*, obtainable for \$1.50 from the Washington office of the APA.

Early Publication. Authors may secure the early publication of an acceptable article in the first issue to go to press after its acceptance by arranging for the article to be printed as extra pages, increasing the number of pages that subscribers receive. Information about charges for early publication may be obtained from the Managing Editor.

National Auxiliary Publications Service. To save printing costs, materials supplementary to articles may be deposited with the NAPS. To obtain copies, order by Document Number, as cited in the reference footnote, from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, New York 10001. The indicated fee must be remitted in advance.

Abstracts. Manuscripts must be accompanied by an abstract of 100–120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or they may be obtained from the Editor or from the APA Central Office.

Copyrighted material. The following policies govern reprinting or other forms of reproduction of APA-copyrighted journal materials. Permission by the American Psychological Association must be requested for reprinting all tables, figures, and text of more than 500 words (total from one article) in length. There will be a charge of \$5.00 per page for such permission except when the requester is the author of the article to be reprinted, or the article is to be reproduced in a limited number of copies solely for instructional purposes.

Reprints. Authors receive 50 reprints gratis. Additional quantities may be ordered when returning proofs.

MANIPULATION OF INDICATIONS OF DISTURBED THINKING IN NORMAL SUBJECTS¹

LEONARD S. LEVITZ² AND LEONARD P. ULLMANN

University of Illinois

A behavioral model of schizophrenia would predict that "normal" people emit indications of disturbed thinking (IDT) and that the rate of emitted IDT will be affected by environmental manipulations. Two 2 × 2 (presence or absence of reinforcement; use or nonuse of instructions) experiments were conducted with college students. In one, Holtzman inkblots (HIT) served as the pretest-posttest measure and a word association test (WAT) was used for training; in the second, the WAT was the pretest-posttest measure and HIT was used for training. Results indicated that (a) both reinforcement and instructions affected IDT rates on both HIT and WAT during training; (b) training results generalized to the pretest-posttest task; and (c) compared to reinforcement, instruction led to a more rapid but less durable increase in IDT.

In contrast to other theories of schizophrenia (e.g., Buss, 1966, pp. 237-390; Millon, 1967; Page, 1966), a behavioral approach proposes that the designation of certain people or their behavior as schizophrenic is the result of differential social reinforcement. Among data cited by Ullmann and Krasner (1969, esp. pp. 186-217 and 377-413) leading to such a formulation are the responsiveness of patients to different social situations (Zarlock, 1966), experimental conditions (Little, 1966; Meichenbaum, 1966; Ullmann, Forsman, Kenny, McInnis, Unikel, & Zeisset, 1965), placebos (Haas, Fink, & Hartfelder, 1963), and direct therapeutic maneuvers (Atthowe & Krasner, 1968; Ayllon & Azrin, 1965; King, Armitage, & Tilton, 1960).

A key question for any theory of schizophrenia is the specification of conditions that lead certain people to be hospitalized. Sociologists such as Scheff (1966) have pointed out that studies of the prevalence of mental illness indicate that there are far more people in the community who are "treatable" (i.e.,

acting in a manner considered disturbed) than there are people who are "treated" (i.e., hospitalized or under psychiatric care). Scheff further notes that the tendency for psychiatrists when making decisions about hospital admission or court commitment is the "safe" one of designating the person psychotic. A related consideration is the unsatisfactory reliability of psychiatric diagnosis, especially when the individual case is the focus. In this situation, very fine and even trivial behavioral manifestations may be used to justify a designation of a symptom, for example, disorganization of thinking, which in turn justifies a diagnosis, for example, schizophrenia, and therefore hospitalization of the person as the solution to a social problem (Glasscote, Cumming, Hammersley, Ozarin, & Smith, 1966; Smith, Pumphrey, & Hall, 1963; Szasz, 1965).

The person designated as schizophrenic is neither passive nor inert. Hospitalized psychiatric patients engage in "impression management" (Braginsky & Braginsky, 1967; Braginsky, Grosse, & Ring, 1966) and, in turn, are affected by their hospital experience (McInnis & Ullmann, 1967). In a similar fashion, rather than the designation as schizophrenic being relatively fortuitous, it is possible to hypothesize that nonhospitalized people emit a low base rate of behaviors that might be called schizophrenic. If some valued change (avoidance or escape from unpleasant

¹ This study is based on a thesis submitted by the first author, under the direction of the second author, in partial fulfillment of requirements for the master's degree at the University of Illinois. The authors wish to thank J. S. Wiggins for his comments and Meryl Bernstein for her assistance with reliability.

² Requests for reprints should be sent to Leonard S. Levitz, 312 West Springfield, Urbana, Illinois 61801.

stimuli or an increase of pleasant stimuli such as acceptance and understanding) occurs after emission of such "disturbed" behavior, the disturbed behavior will be emitted more frequently.

The purpose of this paper is to present data consistent with this behavioral formulation. The Ss used were students attending college. The "schizophrenic" behavior selected was disturbance of language and thought, a variable a number of theorists (e.g., Meehl, 1962) have made central to their formulations of schizophrenia.

Two instruments that are frequently used to measure disturbance of language and thought are the Kent-Rosanoff Word Association Test (WAT) and the Holtzman Inkblot Test (HIT). The WAT provides a definition of disturbed thought based on response frequencies in the normal population, while certain indexes of the HIT, especially those of Pathognomic Verbalization, Integration, and Form Appropriateness, have been found to be significantly associated with scores of the process-reactive distinction (Steffy & Becker, 1961; Ullmann & Eck, 1965) and, by implication, with the severity of the disorder.

A number of recent investigations, some of which were noted above, have indicated that the language behavior of hospitalized schizophrenics may be manipulated directly through the use of operant conditioning techniques. Ullmann, Krasner, and Edinger (1964), Panek (1966), and Thompson (1967) all found that schizophrenics emit an increased rate of common associates to the WAT when reinforced for such behavior. Thompson (1967) and Meichenbaum (1967) have found that another effective method of increasing the emission by schizophrenics of common associates and abstract proverb interpretations, respectively, was the provision of specific instructions about the task such as asking the patient to "act as most people would."

Palermo and Jenkins (1964), in their tabulations of response frequencies to the WAT, have shown that across all age levels there are normal people who give responses that are relatively rare or unusual. Holtzman, Thorpe, Swartz, and Herron (1961) found that college students emitted a low number

of HIT responses that were scorable as Pathognomic Verbalization, poor Integration, or poor Form Appropriateness. Morgan (1968) replicated this finding using a group of nonhospitalized applicants to a psychological clinic. As such, a base rate of indications of disturbed thinking (IDT) as measured by the WAT and HIT may be found in the general population. The present investigation, therefore, centered on whether instructions and response-contingent reinforcement would significantly increase the emission of IDT as measured by the WAT and HIT. If the experimental manipulations were effective, an additional line of investigation would be whether such altered behavior generalized to a different but related task.

METHOD

Instruments

Word Association Test (WAT). In this investigation an IDT was defined as a response that was given by less than 5% of the normative sample of 500 male college students tested by Palermo and Jenkins (1964). Because of the nature of the definition and of the normative sample, it was necessary to eliminate four words from the original word list.³

Holtzman Inkblot Test (HIT). On the basis of pilot work, 48 inkblots were selected from Forms A and B to maximize the base rate of disturbed responses found among college students.⁴ A response was scored as an IDT if the response could be scored positive on any one of three variables: Pathognomic Verbalization, Integration, and Form Appropriateness. The categories and standards used were those presented by Holtzman et al. (1961, pp. 45-66).

Treatments

Upon entering the testing room, each S was given the following general instructions:

I am sure you have heard about psychologists being interested in inkblot cards and word associations. The purpose of today's meeting is to obtain such responses. To each item that I present, I want you to answer as quickly as possible. In the word association part, I will orally present a common word to which I want you to tell me the

³ The most frequent associate of "city" was "Minneapolis" due to the locale of the normative sample; "high" was eliminated because of its possible confusion with its homonym; and the common associates of "trouble" and "house" accounted for less than 40% of the total responses of the sample.

⁴ For a list of the HIT cards used see Levitz (1968).

first word that comes to you. When I give you an inkblot card, I want you to tell me what the card might look like or what it might represent. You may see several things in each inkblot but try to give me only one response.

Group R (reinforcement only). After a base-rate period, and during the training period only, Ss in this group received positive reinforcement contingent on the production of indications of disturbed thinking, IDT. The reinforcing stimulus consisted of a show of attention and a variety of encouraging remarks.

Group I (instructions only). After a base-rate period, Ss in this group received the following specific instructions:

I think your responses would be more spontaneous if you let yourself go: your first thoughts are best. In general, this can be done by acting more like the common stereotype of a mentally ill person. In other words, act like a person who doesn't care what impression he makes on anyone but says whatever he feels like saying.

No reinforcement was given and the instructions were not repeated.

Group RI (reinforcement and instructions). After a base-rate period, Ss received the same specific instructions as Group I. During the training period only, these Ss received positive reinforcement contingent on the emission of indications of disturbed thinking, IDT.

Group C (control). The Ss were given the introductory instructions but no further information, reinforcement, or feedback.

Procedures

This study consisted of two experiments alike in procedures and treatments but differing in the ways the dependent variables were used.

Prior to testing, each S was randomly assigned to one of the four treatment conditions described above and to one of the two experiments.⁵ In each experiment the variables were reinforcement or no reinforcement and instructions or no instructions. In the first experiment, the HIT was used as a pretest and a posttest measure and the WAT was used as the training material. In the second experiment, the pretest and the posttest material was the WAT and the training material was the HIT.

Experiment I. The 48 Ss, 12 in each of the four treatment groups in this experiment, received a pretest of 12 Holtzman inkblots, then a base-rate period of 24 words. There followed a training period of

72 word associations, which was later divided into three subperiods of 24 words each for the purpose of statistical analysis. The training period was followed by a posttest consisting of 12 Holtzman inkblots. Thus, the WAT was used to test whether emission of IDT could be increased, and the HIT provided the measure of response generalization.

Experiment II. The 48 Ss, 12 in each of the four treatment groups in this experiment, received a pretest of association to 24 words and a base-rate period of 12 Holtzman inkblot cards. The training period consisted of 36 inkblots, which was later divided for statistical analysis into three subperiods of 12 inkblots each. The training period was followed by a posttest of association to 24 words.

In Experiment I, the words used as the base-rate measure and the inkblots used in the pretest were identical to the words used in the pretest and the inkblots used in the base-rate period of Experiment II. Further, in Experiment I, the words used in the last training period and the inkblots used in the posttest were identical to the words used in the posttest and the inkblots used in the last training period of Experiment II.

RESULTS

Reliability of Dependent Measures

To determine whether the HIT provided a reliable measure of IDT, Ss' responses were scored by two independent raters: the senior author and an undergraduate student without training in clinical psychology. Prior to being scored, each response was numbered and transcribed to an index card so that the identity of the treatment or test period would be unknown to the raters.

The interrater reliability coefficients, each based on responses by 48 Ss to 12 inkblots were as follows: pretest, .64; posttest, .57; base-rate period, .51; Training Period 1, .66; Training Period 2, .62; Training Period 3, .75. All these correlations were significant beyond the .001 level.

Main Experimental Effects: Experiment I

Before assessing the effects of the experimental treatments on the production of indications of disturbed thinking as measured by the WAT, it was determined that during the base-rate period there were no significant group differences ($F = 2.02$, $df = 3/44$).

In this investigation, it was not only the final level of disturbed thinking that was of interest but also the trend of acquisition that each experimental manipulation produced. The statistical model best suited for deter-

⁵ Of the 96 male Ss in this study, 56 were paid volunteers and 40 were obtained from the introductory psychology course S pool. The Ss were randomly assigned to one of the treatment groups with one restriction: in each group there were 7 Ss from the paid volunteer group and 5 Ss from the S pool.

TABLE 1

ANALYSIS OF VARIANCE OF THE MAIN EXPERIMENTAL
EFFECTS OF EXPERIMENT I: WORD
ASSOCIATION TEST SCORES

Source	<i>df</i>	<i>MS</i>	<i>F</i>
Experimental groups	3	384	8.17*
Experimental periods	3	221	34.53*
Groups \times Periods	9	26	4.06*
Correlated error	44	47	
Uncorrelated error	132	6.4	

* $p < .001$.

mining these effects was the two-way repeated measures ANOVA (Lohdahl, 1967). The results of this analysis are presented in Table 1. Group, period, and the Group \times Experimental Period interaction effects were all significant at the .001 level. By examination of Figure 1, and by a series of post hoc comparisons, the nature of these differences was determined.

In Figure 1 it may be seen that immediately after the special instructions were presented to Group I, the number of indications of disturbed thinking increased sharply. However, this higher rate was not maintained during the succeeding periods, and by the last training period the rate had returned to the base-rate level. The difference between the base-rate period and Training Period 1 was significant at the .01 level, while the difference between the base rate and Training Period 3 was not significant.

Group RI, to which the same instructions were given but which also received positive reinforcement, showed the same sharp increase as Group I. However, this higher rate was maintained to the end of the training periods. Thus, the differences between the base-rate period and Training Periods 1 and 3 were significant at the .01 level.

Receiving positive reinforcement alone, Group R showed a steadily increasing rate of IDT. The differences between the base rate and Training Periods 2 and 3 were significant at the .01 level.

Main Experimental Effects: Experiment II

It was determined that no significant group differences existed during the base-rate period ($F = 2.38$, $df = 3/44$). The results of the repeated measures ANOVA are presented in

Table 2. The differences among experimental groups and among the test periods were significant at the .01 level. The interaction was not significant.

Immediately after the special instructions were given, as may be seen in Figure 2, Group I evidenced a rapid increment in IDT. During Training Period 2, as in Experiment I, this rate declined. However, during the last training period, a second sharp increment appeared. The differences between the base-rate period and Training Periods 1 and 3 were significant at the .01 level.

Group RI showed the same sharp increment after the instructions were presented. This increment was maintained through Training Period 2, but during the last training period the rate decreased. The differences between the base rate and Training Periods 1, 2, and 3 were significant at the .01 level.

Group R showed some increment in the rate of IDT, although not to the extent found on the WAT. Still, the difference between the base rate and Training Period 3 was significant at the .01 level.

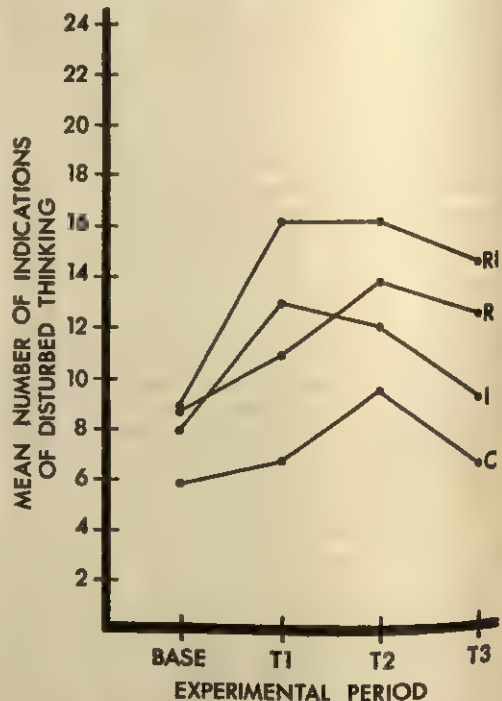


FIG. 1. Experiment I: Main experimental effects (Word Association Test scores).

TABLE 2

ANALYSIS OF VARIANCE OF THE MAIN EXPERIMENTAL EFFECTS OF EXPERIMENT II: HOLTZMAN INKBLLOT TEST SCORES

Source	df	MS	F
Experimental groups	3	26.3	4.61*
Experimental periods	3	23.3	9.71**
Groups \times Periods	9	2.9	1.20
Correlated error	44	5.7	
Uncorrelated error	132	2.4	

* $p < .01$.

** $p < .001$.

Generalization of Experimental Effects: Experiment I

Upon undertaking this investigation, the HIT was believed to measure approximately the same type of behavior (in the present paper, called indication of disturbed thinking, IDT), as did the WAT. The authors were unable to find a study reporting on the correlation between these measures, especially one using normal Ss. Unless a significant positive relationship between the two measures did exist, it would not have been reasonable to expect the results of the training periods to generalize to the second variable. Furthermore, a stringent test of generalization would be to show that the degree of change due to an experimental manipulation on one measure was associated with change on the second measure when that manipulation was no longer being applied.

A correlation coefficient of .46 ($p < .001$, one-tailed) was found between the base-rate period scores of the WAT and the pretest scores of the HIT. Thus, a moderate degree of association existed between the two variables used as measures of disturbed thinking. It was also determined that no significant group differences existed during the pretest period ($F = .13$, $df = 3/44$).

To determine differences among groups from the pretest to the posttest periods, the repeated measures ANOVA was again used. The results of the analysis are presented in Table 3. The only significant finding was a pretest to posttest trial effect across all experimental groups at the .05 level. By examination of Figure 3, and by post hoc compari-

sons, it was determined that the difference between Groups R and I combined, and Group C was significant at the .05 level. All individual comparisons were not significant.

As a final test of generalization, the relationship between the degree of IDT change exhibited on the WAT during training and the degree of IDT change from pretest to posttest HIT was determined. The resulting correlation of +.24 was in the predicted direction and significant at the .05 level ($CR = 1.64$), one-tail test. Considering the rater reliability of the HIT, the moderate degree of association between the two measures, and such factors as time lapse differences, this was not an unimportant finding.

Generalization of Experimental Effects: Experiment II

The correlation coefficient between the base-rate period of the HIT and the pretest period of WAT was found to be .34, significant at the .02 level, one-tail test. Thus, the corresponding finding in Experiment I was replicated. No significant group differ-

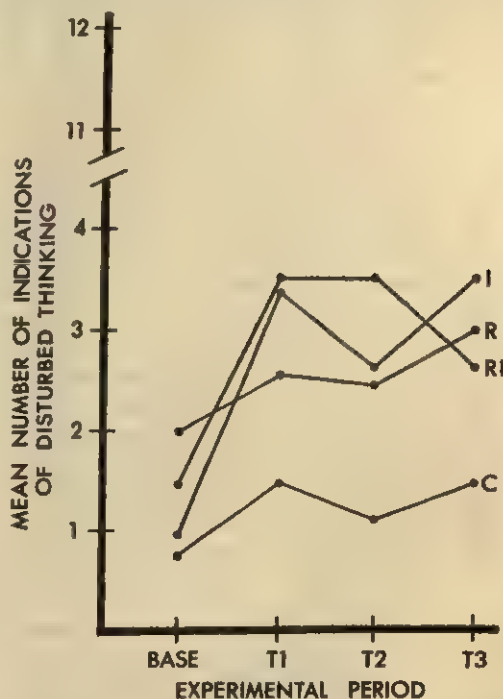


FIG. 2. Experiment II: Main experimental effects (Holtzman Inkblot Test scores).

TABLE 3

ANALYSIS OF VARIANCE OF GENERALIZATION EFFECTS
OF EXPERIMENT I: HOLTZMAN
INKBLOT TEST SCORES

Source	<i>df</i>	<i>MS</i>	<i>F</i>
Experimental groups	3	4.67	1.67
Pre-post periods	1	8.00	4.82*
Groups \times Periods	3	3.00	1.80
Correlated error	44	2.80	
Uncorrelated error	44	1.66	

* $p < .05$.

ences were found during the pretest period ($F = 1.16$, $df = 3/44$).

A summary of the results of the analysis of variance for changes from pretest to posttest is presented in Table 4. The pretest to posttest trials effect was highly significant, as was the Experimental Group \times Test Period interaction. Inspection of Figure 4 reveals that Group I and Group RI showed a sharp pretest to posttest increment as compared with Groups C and R. Remember that during the training periods on the HIT it was Group I that showed the greatest final rate of IDT. The pretest to posttest change of

TABLE 4

ANALYSIS OF VARIANCE OF GENERALIZATION EFFECTS
OF EXPERIMENT II: WORD
ASSOCIATION TEST SCORES

Source	<i>df</i>	<i>MS</i>	<i>F</i>
Experimental groups	3	40.33	2.20
Pre-post periods	1	114.00	30.24*
Groups \times Periods	3	27.00	7.16*
Correlated error	44	18.30	
Uncorrelated error	44	3.77	

* $p < .001$.

Group I was found to be significant at the .05 level.

An analysis of change scores identical to that described in Experiment I resulted in a correlation coefficient of $+.06$ which was in the predicted direction, but not statistically significant.

DISCUSSION

Both experiments in this investigation indicated that the emission by normals (non-hospitalized college students) of indications of disturbed thinking (IDT) may be significantly influenced by environmental manipulations, such as instructions and reinforcement.

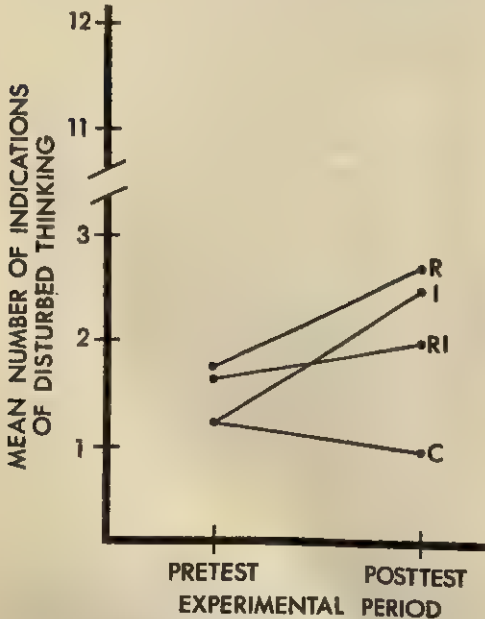


FIG. 3. Experiment I: Generalization effects (Holtzman Inkblot Test scores).

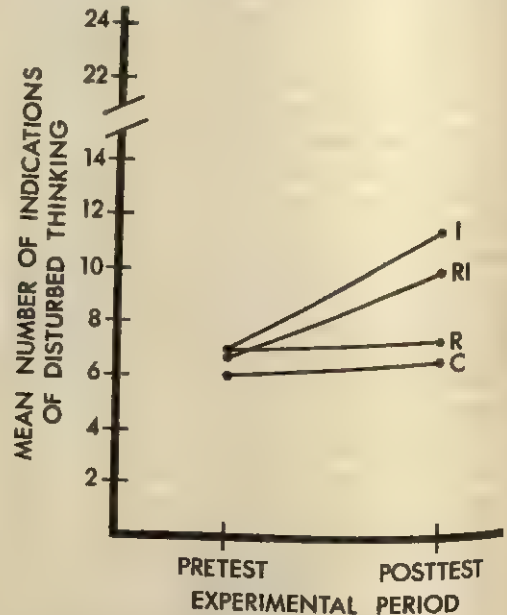


FIG. 4. Experiment II: Generalization effects (Word Association Test scores).

On both experiments, the effects of such training influenced emission of IDT on a different, but correlated task that was also a measure of IDT. These findings are consistent with, and offer an empirical demonstration of, some of the aspects of the behavioral approach to schizophrenia. These data are *not* proof of the behavioral formulation; they do not prove that this is what necessarily happens to people who are called schizophrenic and hospitalized. An experiment which accomplished this would be patently unethical. Further, even if one did shape an experimental group to the point of incarceration, such a procedure would not logically rule out that other factors might also eventuate in people being placed in psychiatric hospitals. To repeat, the data presented are consistent with, but not proof of, the behavioral approach.

To some extent, the present data are inconsistent with other theories of the etiology of schizophrenia. There is no reason to believe that the students underwent regressions or suffered from excessive variations of motivation, attention, social censure, stimulus input, or biochemical substances. While the present work does not rule out the conceivable importance of these latter factors, it does indicate that instructions and reinforcement may also increase the emission of behavior called schizophrenic.

There seem to be two ways in which the present investigation might be attacked. The first is to take the position that it is not the behavior but its antecedents that determine whether the behavior is pathological. This indeed is the position taken by Wegrocki (1939):

The patient's delusion is an internal resolution of a problem; it is his way of meeting the intolerable situation. That is why it is abnormal. It represents a spontaneous protective device of the personality, something which is not learned [p. 169].

Aside from the difficulties involved in formulating behavior as spontaneous and unlearned, the empirical material necessary for a distinction between clinical (or real) and experimental (or unreal) pathology is far from readily available. The relevance of the present work to a behavioral formulation of the etiol-

ogy of schizophrenia could be negated if one denied that the HIT and WAT are measures of IDT or that IDT plays an important role in the designation of a person as schizophrenic. Paradoxically, the formulation of schizophrenia that could most comfortably make these assumptions is the behavioral approach which considers both the designator and recipient of the diagnosis as playing social roles.

At a more directly empirical level, when the present results are compared with studies that used the same manipulations with different tasks, the effects are found to be consistent. Levy (1967), finding that instructions led to an abrupt change in size constancy perception, concluded that verbal reinforcement was merely an "inefficient way of telling the subject what to do." Ayllon and Azrin (1964) found, however, that instructions alone did not have enduring effects: only when they were used in conjunction with contingent reinforcement was a high rate of performance maintained. In the present study, by suggesting that students act like mentally ill people, the need for gradual shaping was eliminated: a discriminative stimulus was provided for a behavior that was already available. Compared to reinforcement, instructions were followed by a more abrupt increase in IDT.

Only under specific conditions, however, should reinforcement be considered "a poor form of instruction": when the target behavior is clearly in the person's repertoire. It is likely that among people who do not possess a clear picture of how the mentally ill should behave (whether this conception is the correct one in the culture or not), such as very young children, the effects of instruction and reinforcement would be quite different. Further, as demonstrated in both experiments of this investigation and by Ayllon and Azrin (1964), the conditions encouraging the emission of a behavior may be different from those needed to maintain it. At a more general level, one might say that if the person does not know the role, he cannot "manage the impression," and, in turn, that if the behavior is not (a) recognized as a particular social role, (b) designated as such by the appropri-

ate societal label, or (c) followed by valued consequences, its level of emission will not increase. This view is similar to that of Becker (1963) on marijuana smoking and of Benedict (1934) when she noted: "Even in trance the individual holds strictly to the rules and expectations of his culture, and his experience is as locally patterned as a marriage rite or an economic exchange [p. 77]."

REFERENCES

- ATTHOWE, J., & KRASNER, L. Preliminary report on the application of contingent reinforcement procedures (token economy) on a "chronic" psychiatric ward. *Journal of Abnormal Psychology*, 1968, 73, 37-43.
- AYLLON, T., & AZRIN, N. H. Reinforcement and instructions with mental patients. *Journal of the Experimental Analysis of Behavior*, 1964, 7, 327-331.
- AYLLON, T., & AZRIN, N. H. The measurement and reinforcement of behavior of psychotics. *Journal of the Experimental Analysis of Behavior*, 1965, 8, 357-383.
- BECKER, H. S. *Outsiders*. New York: Free Press, 1963.
- BENEDICT, R. Anthropology and the abnormal. *Journal of General Psychology*, 1934, 10, 59-82.
- BRAGINSKY, B. M., & BRAGINSKY, D. D. Schizophrenic patients in the psychiatric interview: An experimental study of their effectiveness at manipulation. *Journal of Consulting Psychology*, 1967, 31, 543-547.
- BRAGINSKY, B. M., GROSSE, M., & RING, K. Controlling outcomes through impression management: An experimental study of the manipulative tactics of mental patients. *Journal of Consulting Psychology*, 1966, 30, 295-300.
- BUSS, A. *Psychopathology*. New York: Wiley, 1966.
- GLASSCOTE, R. M., CUMMING, E., HAMMERSLEY, D. W., OZARIN, L. D., & SMITH, L. H. *The psychiatric emergency*. Washington, D. C.: Joint Information Service, 1966.
- HAAS, H., FINK, H., & HARTFELDER, G. The placebo problem. *Psychopharmacology Service Center Bulletin*, 1963, July, 1-65.
- HOLTZMAN, W. H., THORPE, J. S., SWARTZ, J. D., & HERRON, W. H. *Inkblot perception and personality*. Austin: University of Texas Press, 1961.
- KING, G. F., ARMITAGE, S. G., & TILTON, J. R. A therapeutic approach to schizophrenics of extreme pathology: An operant-interpersonal method. *Journal of Abnormal and Social Psychology*, 1960, 61, 276-286.
- LEVITZ, L. S. The manipulation of indications of disturbed thinking in normal subjects. Unpublished master's thesis, University of Illinois, 1968.
- LEVY, H. Effects of verbal reinforcement and instruction on size constancy. *Canadian Journal of Psychology*, 1967, 21, 81-91.
- LOHDAHL, S. *Modern statistics for behavioral sciences*. New York: Ronald Press, 1967.
- LITTLE, L. K. Effects of the interpersonal interaction on abstract thinking performance in schizophrenics. *Journal of Consulting Psychology*, 1966, 30, 158-164.
- MCINNIS, T. L., & ULLMANN, L. P. Positive and negative reinforcement with short- and long-term hospitalized schizophrenics in a probability learning situation. *Journal of Abnormal Psychology*, 1967, 72, 157-162.
- MEEHL, P. E. Schizotaxia, schizotypy, schizophrenia. *American Psychologist*, 1962, 17, 827-838.
- MEICHENBAUM, D. H. Effects of social reinforcement on the level of abstraction in schizophrenics. *Journal of Abnormal and Social Psychology*, 1966, 71, 354-362.
- MEICHENBAUM, D. H. Effects of social reinforcement on the level of abstraction in schizophrenics. *Dissertation Abstracts*, 1967, 27, (7B0) 2514-2515.
- MILLON, T. (Ed.) *Theories of psychopathology*. Philadelphia: Saunders, 1967.
- MORGAN, A. Some age norms obtained for the Holtzman Inkblot Technique administered in a clinical setting. *Journal of Projective Techniques and Personality Assessment*, 1968, 32, 165-172.
- PAGE, J. D. (Ed.) *Approaches to psychopathology*. New York: Columbia University Press, 1966.
- PALERMO, D. S., & JENKINS, J. J. *Word association norms*. Minneapolis: University of Minnesota Press, 1964.
- PANEK, D. M. Word association learning by chronic schizophrenics on a token economy ward under conditions of reward and punishment. *Dissertation Abstracts*, 1966, 27, (6B) 2143-2144.
- SCHEFF, T. J. *Being mentally ill*. Chicago: Aldine, 1966.
- SMITH, K., PUMPHREY, M. W., & HALL, J. C. The "last straw": The decisive incident resulting in the request for hospitalization in 100 schizophrenic patients. *American Journal of Psychiatry*, 1963, 120, 228-232.
- STEFFY, R. A., & BECKER, W. C. Measurement of the severity of disorder in schizophrenia by means of the Holtzman Inkblot Test. *Journal of Consulting Psychology*, 1961, 25, 555.
- SZASZ, T. S. *Psychiatric justice*. New York: Macmillan, 1965.
- THOMPSON, E. E. Effectiveness of different techniques in training schizophrenic patients to give common associations. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, May 1967.
- ULLMANN, L. P., & ECK, R. Inkblot perception and the process-reactive distinction. *Journal of Clinical Psychology*, 1965, 21, 311-313.
- ULLMANN, L. P., FORSMAN, R. G., KENNY, J. W., MCINNIS, T. L., UNIKEL, I. P., & ZEISSET, R. M. Selective reinforcement of schizophrenics' interview responses. *Behaviour Research and Therapy*, 1965, 2, 205-212.

- ULLMANN, L. P., & KRASNER, L. *A psychological approach to abnormal behavior*. Englewood Cliffs, N. J.: Prentice-Hall, 1969.
- ULLMANN, L. P., KRASNER, L., & EDINGER, R. L. Verbal conditioning of common associations in long-term schizophrenic patients. *Behaviour Research and Therapy*, 1964, 2, 15-18.
- WEGROCKI, H. J. A critique of cultural and statistical concepts of abnormality. *Journal of Abnormal and Social Psychology*, 1939, 34, 166-178.
- ZARLOCK, S. P. Social expectations, language, and schizophrenia. *Journal of Humanistic Psychology*, 1966, 6, 68-74.

(Received October 14, 1969)

EDITOR'S NOTE

The following two discussions by Gerald Rosenbaum and Loren J. Chapman were prepared at the request of the Editor.

SCHIZOPHRENIA AS A "PUT-ON"

GERALD ROSENBAUM¹

Wayne State University

Levitz and Ullmann found that instructions and social reinforcement enable college students to enact the "mentally ill" role on two common indicators of schizophrenic thinking and generalized their findings as a paradigm for how schizophrenic thinking is shaped. Evidence is cited that schizophrenics are unresponsive to social instructions and verbal reinforcement, and differ from normals in being unable to turn off disturbed thinking. The importance of determining reinforcement contingencies and the mechanisms responsible for schizophrenic thinking in order to modify central schizophrenic symptoms is emphasized.

The notion that the schizophrenic person fails to learn appropriate social roles and patterns of consensually validated language has been well enunciated by Cameron (1944) and Sullivan (1947) in social learning theories of schizophrenia. These theories have attempted to account for the pervasive social disarticulation of the schizophrenic in the face of rather consistent social nonreinforcement or punishment for the emission of such schizophrenic symptoms as thinking disorder, hallucinations, delusions, or social withdrawal. Other learning theories have attempted to derive the failure of schizophrenics to respond appropriately with social reinforcements from socially learned mechanisms, such as fear of social censure (Rodnick & Garnezy, 1957), a predisposition to disorganize under the stress of high anxiety (Broen & Storms, 1966; Mednick, 1968), or an acquired deficit in adient social motives (Rosenbaum, Mackavey, & Grisell, 1957), and in maintaining a goal-oriented major task set (Shakow, 1962).

A more recent "sociopsychological formulation of schizophrenia" which appears to take exception to all extant psychological positions has been proposed by Ullmann and Krasner (1969). They suggest that enacting the "mentally ill" schizophrenic role may be reinforced by the short-term advantages of a deviant social role (e.g., avoiding army induction, being disturbing to others, and other means of avoiding unpleasant and increasing pleasant stimuli). When this deviant behavior leads to the assignment of the schizophrenic diagnostic label, further training in the role will ensue by modeling and reinforcements in the psychiatric interview and in the hospital complex. The maintenance of appropri-

ate schizophrenic symptoms is assumed to be attributable to their selective reinforcement by the culture of the large psychiatric hospital. Typical of the evidence cited in support of this formulation are experimental studies, such as Levitz and Ullmann (1969), which are aimed at demonstrating that disturbed thinking or other primary behavioral indications of schizophrenia may be shaped by verbal operant conditioning or instructions.

The Levitz and Ullmann study is certainly a rather neat demonstration that college students can learn to enact the mentally ill role on at least two common clinical indicators of schizophrenic thinking. It is also patently clear that these Ss can very readily follow instructions to stop "putting-on" like schizophrenia and can easily turn off these indications of disturbed thinking. Unfortunately, the same is not true for schizophrenics.

While Ullmann and Krasner may regard schizophrenia as a "put-on" and Szasz (1965) may perceive schizophrenic hospitalization as a form of social injustice perpetrated by a system which punishes deviance, the tragedy for the schizophrenic individual is his inability to turn off his disturbed thinking and to opt for himself. The social psychological critiques of hospital practices by writers like Goffman (1961) and Dunham and Weinberg (1960), and the development of token economies (Ayllon & Azrin, 1965) have contributed positively to the growth of more active programs for returning schizophrenic persons to the community. Nevertheless, the peculiar learning history and genetic, neurological, perinatal, or biochemical predispositions of the schizophrenic appear to seriously interfere with his response to well-conceived social reinforcements even in this era of community participation in the treatment program through the unit system.

¹Requests for reprints should be sent to Gerald Rosenbaum, Department of Psychology, Wayne State University, Detroit, Michigan 48202.

In view of these considerations and the intractability of schizophrenic symptoms to durable modification, it is difficult to understand how careful investigators like Levitz and Ullmann (1969) can seriously generalize their results to an approximation of "what necessarily happens to people who are called schizophrenic and hospitalized [p. 639]" or to a paradigm enabling one to "shape an experimental group to the point of incarceration [p. 639]." They found that instructions, verbal reinforcement, and preferably their combination produced statistically significant increases in the frequency of uncommon word associations and disturbed Holtzman Inkblot responses emitted by college students, as well as some intertask transfer. The combined experiments permit comparisons of the different training procedures on both tasks and generate the conclusion that instructions lead to more rapid acquisition but less durable increases in these responses. The study demonstrates that college students can simulate the mentally ill role well enough to produce the relevant indications of disturbed thinking under the conditions. Despite Levitz and Ullmann's concern for the unethical implications of unduly shaping disturbed thinking in their Ss, their data indicate that their Ss were in no danger of becoming schizophrenic as their emission of these responses was primarily under the control of self-instructions to play a deviant role on these tasks. With the resumption of normal roles, these Ss might be expected to resume appropriate responses. By contrast, schizophrenics do not resume normal roles, and they do continue to emit disturbed thinking in spite of instructions, reinforcement, and all of the contemporary therapeutic maneuvers that are available for shaping appropriate responses.

It should be noted that the sociopsychological thesis that the schizophrenic is an individual playing the mentally ill role for reasons of short-term self-interest is tantamount to conceptualizing the schizophrenic as a malingerer. The crux of this proposition requires evidence that schizophrenics do respond to self-instructions or social instructions to play more appropriate roles. The present evidence suggests that Ss who could conform to this criterion for malingering are the students in the Levitz and Ullmann study, not schizophrenics. The theoretical problem of schizophrenia is the explanation of the failure of social instructions and reinforcement to appreciably modify symptomatic behavior in approximately 1% of the population. The sociopsychological formulation solves this problem by denying it.

The recent interest in social psychological analysis of schizophrenia appears to be partly attributable to growing enthusiasm for operant analysis of social processes. The particular position of Levitz and Ullmann appears to rest also on certain assumptions concerning learning effects and thought disorder in schizophrenia.

Learning and Schizophrenia

In contrast to mental retardates, diminished learning rate does not appear to be a primary parameter of schizophrenia. Both clinical and experimental data on a number of tasks (Jones, 1961; Lindsley, 1956) indicate that schizophrenics show normal learning functions when reinforcements are appropriate. The fact that schizophrenics show instrumental learning of both symptomatic and more appropriate behaviors is cited as central to the Ullmann and Krasner (1969) thesis that schizophrenia is primarily a "put-on" in which certain individuals are tagged with this pejorative label because their deviant role playing is disturbing to others. Yet Meehl (1962) concludes his penetrating analysis of this problem with the conviction "that schizophrenia, while its content is learned, is fundamentally a neurological disease of genetic origin [p. 837]." Instrumental learning in schizophrenia is not surprising and continues to occur in all organisms with both acquired and physiological deficits as they adapt to their environments.

The treatment issue in schizophrenia is the determination of the conditions for and types of reinforcement that will increase the operant rate of appropriate responses and minimize symptomatic behaviors, regardless of etiology. Work with token economies (Ayllon & Azrin, 1965) has leaned heavily on the use of candy, cigarettes, and other amenities of life as reinforcers, rather than verbal and social approval. Premack (1959) has ingeniously suggested that the opportunity to perform high-frequency behaviors, even symptoms, may be employed as reinforcers in behavior modification. In summarizing the literature, Buss and Lang (1965) note that biologically noxious stimulation is the only contingency that has led to complete elimination of schizophrenic deficit (e.g., Rosenbaum, 1967), while motivating instructions (Cohen, 1956) and positive verbal reinforcement (Cohen & Cohen, 1960) have been least effective in modifying schizophrenic performance.

Thought Disorder in Schizophrenia

As Levitz and Ullmann (1969) suggest, most investigators regard language and thought dis-

order as central to schizophrenic pathology and therefore as the most important behavior to shape. To date, the powerful control and reinforcers used in token economies have been employed to shape improvements in grooming, participation, cooperation, and ward morale (e.g., Atthowe & Krasner, 1968), but there have been no reports of the elimination of thought disorder. The Levitz and Ullmann demonstration that normal Ss can be shaped to make more uncommon associations, coupled with the findings by Ullmann, Krasner, and Edinger (1964) that schizophrenics can be trained to make more common word associations, are empirical contributions toward this objective. The failure of schizophrenics to maintain such experimentally induced improvements in language and of normals to retain trained thought disturbances probably reflects the long-standing differences in language hierarchies and socialization implied by the psychological learning theories of Cameron (1944) and Sullivan (1947). Effective shaping of schizophrenic thinking is likely to require more than the verbal reinforcement of common associates. It seems reasonable to assume that some theoretical formulation of the habit strengths and reinforcement histories of schizophrenic language hierarchies and the mechanisms leading to disturbed verbal behavior would be critical to durable shaping.

Recent work on schizophrenic language by Chapman, Chapman, and Miller (1964) suggests that thought disorder in schizophrenia results from the associative intrusion of strong dominant meaning responses in the mediation of overt language behavior. Cromwell and Dokecki (1968) have lately reviewed the principal theories of schizophrenic language and offered their own disattention interpretation, that schizophrenics are deficient in the ability to withdraw attention from prior stimuli. Meehl (1962) has even suggested that "cognitive slippage" is the result of a central neural integrative deficit which interacts with specific genetic, biochemical, and learning mechanisms. The development of sound programs for shaping disturbed thinking in the varied conditions designated as schizophrenia is likely to use a number of these theoretical formulations. There may even be some schizophrenics who are best understood as a "put-on."

REFERENCES

- ATTOWE, J. M., JR., & KRASNER, L. A preliminary report on the application of contingent reinforcement procedures (token economy) on a "chronic" psychiatric ward. *Journal of Abnormal Psychology*, 1968, 73, 37-43.
- AYLON, T., & AZRIN, N. H. The measurement and reinforcement of behavior of psychotics. *Journal of the Experimental Analysis of Behavior*, 1965, 8, 357-384.
- BROEN, W. E., JR., & STORMS, L. H. Lawful disorganization: The process underlying a schizophrenic syndrome. *Psychological Review*, 1966, 73, 265-279.
- BUSS, A. H., & LANG, P. J. Psychological deficit in schizophrenia: I. Affect, reinforcement, and concept attainment. *Journal of Abnormal Psychology*, 1965, 70, 2-24.
- CAMERON, N. The functional psychoses. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald Press, 1944.
- CHAPMAN, L. J., CHAPMAN, J. P., & MILLER, G. A. A theory of verbal behavior in schizophrenia. In B. A. Maher (Ed.), *Progress in experimental personality research*. Vol. 1. New York: Academic Press, 1964.
- COHEN, B. D. Motivation and performance in schizophrenia. *Journal of Abnormal and Social Psychology*, 1956, 52, 186-190.
- COHEN, E., & COHEN, B. D. Verbal reinforcement in schizophrenia. *Journal of Abnormal and Social Psychology*, 1960, 60, 443-446.
- CROMWELL, R. L., & DOKECKI, P. R. Schizophrenic language: A disattention interpretation. In S. Rosenberg & J. H. Koplin (Eds.), *Developments in applied psycholinguistics research*. New York: Macmillan, 1968.
- DUNHAM, W. H., & WEINBERG, S. K. *The culture of the state mental hospital*. Detroit: Wayne State University Press, 1960.
- GOFFMAN, E. *Asylums*. Garden City, N. Y.: Doubleday, 1961.
- JONES, H. G. Learning and abnormal behavior. In H. J. Eysenck (Ed.), *Handbook of abnormal psychology*. New York: Basic Books, 1961.
- LEVITZ, L. S., & ULLMANN, L. P. Manipulation of indications of disturbed thinking in normal subjects. *Journal of Consulting and Clinical Psychology*, 1969, 33, 633-641.
- LINDSLEY, O. R. Operant conditioning methods applied to research in chronic schizophrenia. *Psychiatric Research Reports*, 1956, 5, 118-153.
- MEDNICK, S. A. A learning theory approach to research in schizophrenia. *Psychological Bulletin*, 1958, 55, 316-327.
- MEEHL, P. E. Schizotaxia, schizotypy, schizophrenia. *American Psychologist*, 1962, 17, 827-838.
- PREMACK, D. Toward empirical behavior laws. I: Positive reinforcement. *Psychological Review*, 1959, 66, 219-233.
- RODNICK, E. H., & GARMEZY, N. An experimental approach to the study of motivation in schizophrenia. *Nebraska Symposium on Motivation*, 1957, 5, 109-184.

- ROSENBAUM, G. Reaction time indices of schizophrenic motivation: A cross cultural replication. *British Journal of Psychiatry*, 1967, 113, 537-541.
- ROSENBAUM, G., MACKAVEY, W. R., & GISELL, J. L. Effects of biological and social motivation on schizophrenic reaction time. *Journal of Abnormal and Social Psychology*, 1957, 54, 364-368.
- SHAKOW, D. Segmental set. *Archives of General Psychiatry*, 1962, 6, 1-17.
- SULLIVAN, H. S. *Conceptions of modern psychiatry*. Washington, D. C.: William A. White Foundation, 1947.
- SZASZ, T. S. *Psychiatric justice*. New York: Macmillan, 1965.
- ULLMANN, L. P., & KRASNER, L. *A psychological approach to abnormal behavior*. Englewood Cliffs, N. J.: Prentice-Hall, 1969.
- ULLMANN, L. P., KRASNER, L., & EDINGER, R. L. Verbal conditioning of common associations in long-term schizophrenic patients. *Behaviour Research and Therapy*, 1964, 2, 15-18.

(Received April 30, 1969)

SCHIZOMIMETIC CONDITIONS AND SCHIZOPHRENIA

LOREN J. CHAPMAN¹

University of Wisconsin

Levitz and Ullmann have added operant reinforcement to the list of allegedly schizomimetic conditions that have been used to generate theories of schizophrenic thought disorder. The diversity of such allegedly schizomimetic conditions casts doubt on the validity of interpreting schizophrenic symptoms on the basis of any one of them. The very power of operant procedures to mimic any behavior, regardless of its origin, argues against invoking its schizomimetic power as evidence concerning the nature of schizophrenia. The potential usefulness of operant methods to alter schizophrenic symptoms need not be predicated on an operant origin of the disorder.

Levitz and Ullmann (1969) have interpreted their findings as consistent with their hypothesis that schizophrenic symptoms are caused by operant reinforcement. The present discussion is focused on their general approach to the problem of schizophrenia, that is, their citing performance of normal Ss in a schizomimetic condition as evidence concerning the origin or nature of schizophrenic thought disorder.

Levitz and Ullmann's study was well designed, and it demonstrated that an operant etiology of schizophrenic symptoms is at least plausible. The investigators concluded, with appropriate caution, that their results are "consistent with" but "not proof of" the operant origin of schizophrenic thought disorder. I agree, except in a matter of relative emphasis on the two sides of the authors' conclusions. There are additional reasons, not discussed by Levitz and Ullmann, for viewing their results as very weak support for the hypothesis that reinforcement history accounts for the origin of schizophrenic thought disorder.

Operant methods are capable of producing almost any conceivable change in voluntary behavior and even, perhaps, some involuntary (autonomic) responses. Verbal responses are simply one more kind of response that can be altered operantly. Could anyone doubt after reviewing the successes of operant workers that one could train college students to rave as paretics, speak inadequately as motor aphasics, or even to quack as ducks? However, would anyone seriously suggest that such a demonstration of the power of operant methods would constitute support for the hypothesis that paretics, aphasics, and ducks fail to speak as normal adult human beings primarily because of un-

fortunate reinforcement histories? The operant method has the power to shape responses that mimic almost any behavior in its natural state, regardless of its origin. Such power rules out any one such demonstration as convincing evidence that behavior of unknown origin occurred primarily because of reinforcement contingencies.

There is a second good reason for caution in viewing the schizomimetic power of reinforcement as support for a primarily operant etiology of schizophrenic thought disorder. Many other schizomimetic conditions have been reported as producing schizophreniclike verbal behavior. Investigators of these conditions have often, like Levitz and Ullmann, argued that the similarity of schizophrenic cognition to that of Ss in some one schizomimetic condition is evidence that the naturally occurring schizophrenic symptom has a cause like that of the behavior elicited in the schizomimetic condition. The very multiplicity of these conditions that produce behavior at least superficially akin to schizophrenic symptoms casts doubt on the validity of the argument for any one of them. The following is at least a partial list of these allegedly schizomimetic conditions that have been invoked as evidence for some particular theory of schizophrenia.

1. LSD, mescaline, and other drugs. Recently, the most widely discussed schizomimetic conditions have been LSD, mescaline, and other drugs. Numerous investigators have reported schizophreniclike cognitive behaviors in response to drugs (Hoffer & Osmond, 1967). These behaviors include disturbed association, pressure of thought, neologisms, blocking, hallucinations and delusions. Several investigators have inferred from such evidence that schizophrenia is caused by substances similar to one or another of these drugs. This hypothesis has stimulated a vast amount of research on the biochemistry of these drugs, and has led to the identification of new

¹ Requests for reprints should be sent to Loren Chapman, Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

schizomimetic drugs (notably adrenochrome, adrenolutin, and sernyl) and, concomitantly, to revised hypotheses of the biochemical basis of schizophrenia.

2. Sensory deprivation. The McGill investigators, who initiated the sensory deprivation research tradition, reported that a prolonged low level of sensory variation produces a deterioration in the ability to think in an organized, goal-directed manner, as well as increased day-dreaming, hallucinations, distortion of body image, and other perceptual aberrations (Bexton, Heron, & Scott, 1954; Heron, Bexton, & Hebb, 1953). Several other investigators have reconfirmed these observations.

Rosenzweig (1959) contended that the sensory deprivation reaction more closely resembles schizophrenia than either the LSD or mescaline psychoses. He suggested that both schizophrenia and the sensory deprivation reaction are caused by a loss of meaning to stimuli. In sensory deprivation, meaning is removed from the environment by the experimental procedures, whereas in schizophrenia, it results from an internal derangement of communication between the patient's "abstract system" and his "affect system."

Gaarder (1963) also cited the similarities between schizophrenia and the cognitive disorders in sensory deprivation and suggested that schizophrenia is the result of lack of input of necessary data about the environment.

3. Sleep deprivation. Many investigators have noted that prolonged sleep deprivation produces cognitive disorganization, hallucinations, and occasional delusions. The suggestion that schizophrenia has its origin in a mechanism similar to that of the symptoms of sleep deprivation has been made by Tyler (1947), Luby, Grisell, Frohman, Lees, Cohen, and Gottlieb (1962), and West, Janszen, Lester, and Cornelisoon (1962). Of these writers, Luby et al. were the most specific. They found in sleep deprivation a disturbance of the synthesis of adenosine triphosphate and they proposed a similar defect as the cause of schizophrenia. Bliss, Clark, and West (1959) contended that naturally occurring schizophrenic episodes are often the direct result of sleep deprivation, pointing to similarities in cognition between schizophrenics and sleep-deprived Ss to bolster their contention.

4. Hypnosis. King (1957) gave a proverbs test to normal Ss, both during hypnosis and in the wakeful state. Because he found that hypnosis increased concrete and bizarre answers, he concluded that schizophrenia, like hypnosis, is a manifestation of heightened suggestion.

5. Speeded performance. Flavell and Draguns (1957), following a lead of Schilder (1951), argued that each normal thought passes through a series of stages of development, and that schizophrenic cognition represents prematurely expressed early stages in the development of normal cognition. Flavell and Draguns (1958) offered as supportive evidence of this theory their finding of some similarities between the various idiosyncrasies of schizophrenics' responses on a word association test and those exhibited by normal Ss when required to respond very rapidly.

Phillips and Framo (1954) reached a similar conclusion about schizophrenic cognition by noting a similarity between the responses that schizophrenics typically gave to the Rorschach blots and those of normal adults presented with brief (tachistoscopic) exposures of the Rorschach blots.

6. Distraction. Jung (1919) studied the effects of distraction on the responses of normal Ss in a word association test; his manipulation of distraction consisted of requiring S to mark a paper with a pencil in time with the beat of a metronome. Jung found that distraction more than doubled the production of rhyme responses and more than quadrupled the production of other clang associations. Later, Jung (1936) invoked the similarity of these responses to those typically found in dementia praecox patients as evidence that the responses of the patients are deviant because of insufficient attention.

Hassol, Magaret, and Cameron (1952) found that normal Ss showed an increase of "scattered speech" in telling TAT stories while simultaneously listening to a recording of their own previous story of the same TAT card. "Scattered speech" included imprecise speech, the use of substitute or approximate words or phrases, and interruptions in the sequence of the material. The investigators concluded that the findings supported Cameron's hypothesis that schizophrenic disorganized speech results from the intrusion of personal fantasy material.

7. Relaxed attention. Jung (1936) reviewed at length a study by Stransky, who obtained samples of continuous speech from normal Ss under the condition of "relaxed attention." Stransky gave a stimulus word and instructed his Ss to speak for 1 minute, but not to pay attention to what they said. The Ss showed many features of speech that are commonly observed in schizophrenics, including perseverations, repetitions, sound associations, and condensations of sequences of ideas into brief phrases or into single words (neologisms). Jung interpreted these find-

ings, and his own distraction studies, as proving that dementia praecox is characterized by a diminution of attention.

8. Disruption of perception. Aaronson (1964), using hypnotism, induced a disturbance in depth perception in a single normal *S*, who responded with psychotic ideation and mood changes. The author inferred that disorder in depth perception may produce schizophrenia. Aaronson (1965), again using a single *S*, reported that he produced psychotic ideation, hallucinations, and withdrawal by communicating to the hypnotized *S* that time had stopped. Aaronson concluded that "Schizophrenia seems specifically a response to the elimination of some conventional dimension of perception [p. 81]."

9. Anoxia. Several investigators, including Freeman (1931), have noted schizophreniclike cognitive aberrations that accompany oxygen deprivation. McFarland's (1932) study was especially thorough. He reported that normal young men with reduced oxygen intake manifest several schizophreniclike symptoms, including defective judgment, perseveration, an abandonment of moral inhibitions, and, subsequently, partial or complete amnesia. In more severe oxygen deprivation, hallucinations and delusions were found to appear. Hinsie, Barach, Harris, Brand, and McFarland (1934) pointed to these phenomena as support for the hypothesis that schizophrenic symptoms might reflect deficient oxygen supply to brain cells. They took their hypothesis so seriously that they converted a patients' dormitory into a large oxygen chamber for three months in an attempt to find a treatment for the disorder.

10. Brain damage. Several influential clinicians have suggested that schizophrenic cognition resembles that of brain-damaged patients, and that schizophrenia is probably, therefore, a similar condition. Kraepelin (1919) named the disorder "dementia praecox" because he believed that these patients show a mental enfeeblement resembling in part that found in paresis, senility, and epilepsy. Many other writers have pointed out specific similarities in cognitive deficit, between schizophrenics and brain-damaged patients, and have inferred that schizophrenics also have cerebral organic malfunction. The deficit cited in common has most often been one of concept formation (Babcock, 1933; Bychowski, 1935, 1943). Goldstein (1939) also took this view, but later reversed himself (Goldstein, 1959).

11. Childhood. Many writers, especially psychoanalytic theorists, pointing to the "primary process" in the thinking of schizophrenics and

that of young children have inferred that schizophrenia results from a regression to an early stage of development (Fenichel, 1945).

12. "Primitive" racial development. Another version of the regression hypothesis is that schizophrenics have regressed to an earlier stage of man's development such as that of prehistoric man or members of contemporary "primitive" societies. Storch (1924), Arieti (1955), and Werner (1948) have pointed to a number of similarities of cognitive functioning that they believe support this hypothesis.

13. Dreams and sleep. The ancient observation that dreams resemble madness has been exploited most fully by psychoanalytic theorists who see "primary process" in both. A number of psychoanalysts have inferred that schizophrenic thought, like that of dreams, reflects an abandonment of the regulative and censoring functions of the ego (Freud, 1953; Rosen, 1953).

Despite the current impasse reached by investigators employing the method of schizomimetic conditions, the method is not logically unsound. If two syndromes are truly identical, one must take seriously the hypothesis that they have a similar cause. However, a major difficulty faced by research workers is the crudity of most measures of schizophrenic thought disorder. Apparent similarities may vanish when specific details of two error patterns are considered. Cameron (1938a, 1938b) is well known for making this point in an attempt to refute both the contention that schizophrenic thought is regressive and that it shows organic deterioration. Chapman, Burstein, Day, and Verdone (1961) have also elaborated this argument in reference to the regression hypothesis. Writers too numerous to mention have recently made the same point in relation to the allegedly schizomimetic drugs. It must be noted that Levitz and Ullmann's measure of similarity on the word association test is one of the least specific in the literature, as they merely trained atypical responses, without any scrutiny of the nature of the atypical responses. However, their measure on the Holtzman Inkblot Test was more specific.

A second difficulty faced by research in this area is the crudity of current systems of diagnosis. Almost all investigators of schizomimetic conditions have not attempted to specify what types of schizophrenics are mimicked. If there are meaningful subdisorders within schizophrenia, one might expect the cognitive disorder of some of them, but not others, to be mimicked by one or another schizomimetic condition. Tutko and Spence's (1962) study exemplifies the promise of this approach. They found that on the passive

sorting portion of the Object Sorting Test, process and reactive schizophrenics depart from normal performance by distinctly different error patterns, and that the errors of brain-damaged patients resemble only those of the process schizophrenics.

Levitz and Ullmann have added to the long list of schizomimetic conditions one more condition that produces deviant verbal behavior like that of schizophrenics. One might ask whether anything that is known about reinforcement should lead one to take it more seriously as a basis of schizophrenic thought disorder than the dozen or so other schizomimetic conditions. The contrary seems true. Reinforcement may change behavior in any direction, toward health as well as illness, while the other schizomimetic conditions almost always debilitate cognition. Moreover, some schizomimetic conditions may produce only changes that are identical to those of some variety of schizophrenia.

Many operant workers might contend that the very power of reinforcement to change behavior lends support to the presumption that in the absence of contrary evidence, any behavioral variation is a learned one. This is a hypothesis worthy of investigation. The most appropriate way to test the hypothesis in reference to schizophrenia is probably by means of longitudinal studies of family interaction. Such studies would, unfortunately, be difficult to perform.

The paucity of evidence for an operant origin of schizophrenic symptoms should not, however, discourage workers who seek to modify schizophrenic symptoms by operant methods. Promising first steps in modifying schizophrenic cognition have already been reported by several investigators, as pointed out by Ullmann and Krasner (1969).

REFERENCES

- AARONSON, B. S. Hypnosis, depth perception and schizophrenia. Paper presented at the meeting of the Eastern Psychological Association, Philadelphia, April 1964.
- AARONSON, B. S. Hypnosis, time rate perception, and personality. Paper presented at the meeting of the Eastern Psychological Association, Atlantic City, April 1965.
- ARIEVI, S. *Interpretation of schizophrenia*. New York: Brunner, 1955.
- BARCOCK, H. *Dementia praecox, a psychological study*. Lancaster, Pa.: Science Press, 1933.
- BEXTON, W. H., HERON, W., & SCOTT, T. H. Effects of decreased variation in the sensory environment. *Canadian Journal of Psychology*, 1954, 8, 70-76.
- BLISS, E. L., CLARK, L. D., & WEST, C. D. Studies of sleep deprivation: Relationship to schizophrenia. *A.M.A. Archives of Neurology and Psychiatry*, 1959, 81, 348-359.
- BYCHOWSKI, G. Certain problems of schizophrenia in light of cerebral pathology. (Trans. by W. H. Wegrocki) *Journal of Nervous and Mental Disease*, 1935, 81, 280-298.
- BYCHOWSKI, G. Physiology of schizophrenic thinking. *Journal of Nervous and Mental Disease*, 1943, 98, 368-386.
- CAMERON, N. A study of thinking in senile deterioration and schizophrenic disorganization. *American Journal of Psychology*, 1938, 51, 650-654. (a)
- CAMERON, N. Reasoning, regression, and communication in schizophrenics. *Psychological Monographs*, 1938, 50(1, Whole No. 221). (b)
- CHAPMAN, L. J., BURSTEIN, A. G., DAY, D., & VERDONE, P. Regression and disorders of thought. *Journal of Abnormal and Social Psychology*, 1961, 63, 540-545.
- FENICHEL, O. *The psychoanalytic theory of neurosis*. New York: Norton, 1945.
- FLAVELL, J. H., & DRAGUNS, J. A microgenetic approach to perception and thought. *Psychological Bulletin*, 1957, 54, 197-217.
- FLAVELL, J. H., DRAGUNS, J., FEINBERG, L. D., & BUDIN, W. A microgenetic approach to word association. *Journal of Abnormal and Social Psychology*, 1958, 57, 1-8.
- FREEMAN, W. Psychochemistry; some physiochemical factors in mental disorders. *Journal of the American Medical Association*, 1931, 97, 293-296.
- FREUD, S. The unconscious. In, *Collected papers*. Vol. 4. London: Hogarth Press, 1953.
- GAARDER, K. A conceptual model of schizophrenia. *Archives of General Psychiatry*, 1963, 8, 590-598.
- GOLDSTEIN, K. The significance of special mental tests for diagnosis and prognosis in schizophrenia. *American Journal of Psychiatry*, 1939, 96, 575-587.
- GOLDSTEIN, K. Concerning the concreteness in schizophrenia. *Journal of Abnormal and Social Psychology*, 1959, 59, 146-147.
- HASSOL, L., MAGARET, A., & CAMERON, N. The production of language disorganization through personalized distraction. *Journal of Psychology*, 1952, 33, 289-299.
- HERON, W., BEXTON, W. H., & HEBB, D. O. Cognitive effects of a decreased variation in the sensory environment. *American Psychologist*, 1953, 8, 366. (Abstract)
- HINSIE, L. E., BARACH, A. L., HARRIS, M. M., BRAND, E., & MCFARLAND, R. A. The treatment of dementia praecox by continuous oxygen administration in chambers and oxygen and carbon dioxide inhalations. *Psychiatric Quarterly*, 1934, 8, 34-71.
- HOFFER, A., & OSMOND, H. *The hallucinogens*. New York: Academic Press, 1967.
- JUNG, C. G. *Studies in word association*. (Trans. by M. D. Eder) New York: Moffat, 1919.
- JUNG, C. G. The psychology of dementia praecox. *Journal of Nervous and Mental Disease*, 1936, Monograph No. 3. (Originally published, 1906).
- KING, P. E. Hypnosis and schizophrenia. *Journal of Nervous and Mental Disease*, 1957, 125, 481-486.

- KRAEPELIN, E. *Dementia praecox and paraphrenia*. (Trans. by R. Mary Barclay) Edinburgh: Livingstone, 1919.
- LEVITZ, L. S., & ULLMANN, L. P. Manipulation of indications of disturbed thinking in normal subjects. *Journal of Consulting and Clinical Psychology*, 1969, 33, 633-641.
- LUBY, E. D., GRISELL, J. L., FROHMAN, C. E., LEES, H., COHEN, B. D., & GOTTLIEB, J. S. Biochemical, psychological and behavioral responses to sleep deprivation. *Annals of the New York Academy of Science*, 1962, 96, 71-79.
- McFARLAND, R. A. The psychological effects of oxygen deprivation (anoxemia) on human behavior. *Archives of Psychology*, 1932, No. 145.
- PHILLIPS, L., & FRAMO, J. L. Developmental theory applied to normal and psychopathological perception. *Journal of Personality*, 1954, 22, 465-474.
- ROSEN, J. N. *Direct analysis*. New York: Grune & Stratton, 1953.
- ROSENZWEIG, N. Sensory deprivation and schizophrenia: Some clinical and theoretical similarities. *American Journal of Psychiatry*, 1959, 116, 326-329.
- SCHILDER, P. On the development of thoughts. In D. Rapaport (Ed.), *Organization and pathology of thought*. New York: Columbia University Press, 1951.
- STORCH, A. The primitive archaic forms of inner experiences and thought in schizophrenia. *Journal of Nervous and Mental Disease*, 1924, Monograph No. 36.
- TUTKO, T. A., & SPENCE, J. T. The performance of process and reactive schizophrenics and brain-injured subjects on a conceptual task. *Journal of Abnormal and Social Psychology*, 1962, 65, 387-394.
- TYLER, D. B. The effect of amphetamine sulfate and some barbiturates on the fatigue produced by prolonged wakefulness. *American Journal of Physiology*, 1947, 150, 253-262.
- ULLMANN, L. P., & KRASNER, L. *A psychological approach to abnormal behavior*. Englewood Cliffs, N. J.: Prentice-Hall, 1969.
- WERNER, H. *Comparative psychology of mental development*. Chicago: Follett, 1948.
- WEST, L. J., JANSZEN, H. H., LESTER, B. K., & CORNELISSEN, F. S., JR. The psychosis of sleep deprivation. *Annals of the New York Academy of Science*, 1962, 96, 66-70.

(Received April 18, 1969)

AN EXPERIMENTAL PROCEDURE FOR THE MODIFICATION OF PSYCHOTIC BEHAVIOR

IVAN TOBY RUTNER¹ AND CHARLES BUGLE

Jacksonville State Hospital, Illinois

A hospitalized patient exhibiting hallucinatory behavior was subjected to a self-monitoring procedure with and without social reinforcement. After 16 days of treatment, hallucinations were extinguished with no remissions reported for approximately 6 months. Other aspects of her behavior also showed improvement.

Of the various new techniques used in modifying behavior, perhaps the most promising are those which emphasize self-control. Self-control can be thought of as a process by which an individual sets up conditions in his environment to bring about specific behaviors in himself (Goldiamond, 1965a). Self-control techniques have been demonstrated to be effective in modifying behaviors ranging from poor studying habits (Sulzer, 1962) to stuttering (Goldiamond, 1965b).

The technique with perhaps one of the simplest applications is self-monitoring. Kanfer (1967) contends that self-control is facilitated by immediate positive feedback. He feels that behavior may be brought under control when accurate feedback is available. By self-monitoring, the individual is provided with this feedback and becomes aware of the effect and consequences of his behavior. Essentially this technique consists of an individual keeping a record of the frequency in which he engages in the behavior to be controlled. This provides the feedback which makes the individual cognizant of this behavior and enables him to bring it under his control. Self-monitoring has demonstrated effectiveness in modifying eating (Stollak, 1966) and smoking (Rutner, 1967) behavior. It has been shown that the effectiveness of self-monitoring may be increased when the procedure is made a "public event" subject to environmental consequences (Rutner, 1967).

The present study was undertaken to determine if self-monitoring with and without social

reinforcement could function as a modifier of psychotic behavior.

PROBLEM

Eva was a 47-year-old female volunteer resident of a state mental hospital diagnosed as schizophrenic. She had been hospitalized for a period of 13 years. Eva sought hospitalization as a result of frequently occurring auditory hallucinations. These hallucinations contributed to a condition of severe depression and anxiety. What was particularly distressing to Eva was the extent to which her behavior was controlled by her "voices." She reported that she was not even permitted access to the bathroom unless it was ordered by the "voices." What was especially intriguing about this problem was Eva's strong desire to resume control of her behavior. She felt that if her "voices" could be brought under control and eliminated she would be able to leave the hospital and function adequately in the community. This opinion was also shared by the hospital staff.

PROCEDURE

As a result of Eva's desire to control her behavior and because of the presence of "adequate personal resources" available to her, a procedure using self-control was selected. These personal resources (Kanfer & Phillips, 1966) include an ability to give reliable self-reports on the basis of self-observation, the capacity of learning some degree of self-control, and a history of responding to the therapist as a positive reinforcer.

Of the self-control techniques that could have been used, self-monitoring was selected because of the relative simplicity of its administration. Eva was instructed to privately record the frequency of her hallucinatory behavior for a 3-day period. She was next instructed to publicly record the daily frequency of hallucinatory behavior on a recording chart which was provided. The chart was strategically placed on the ward, directly behind the nursing station, in full view of everyone on the ward. As a result of the chart placement, anyone on the ward could be aware of Eva's progress at any time. The chart, beside providing immediate feedback to Eva, also functioned as a stimulus for social reinforcement on the part of the nursing staff, experimenters, and

¹ Requests for reprints should be sent to Ivan Toby Rutner, who is now at the University of Manitoba, Department of Psychology, Winnipeg, Canada.

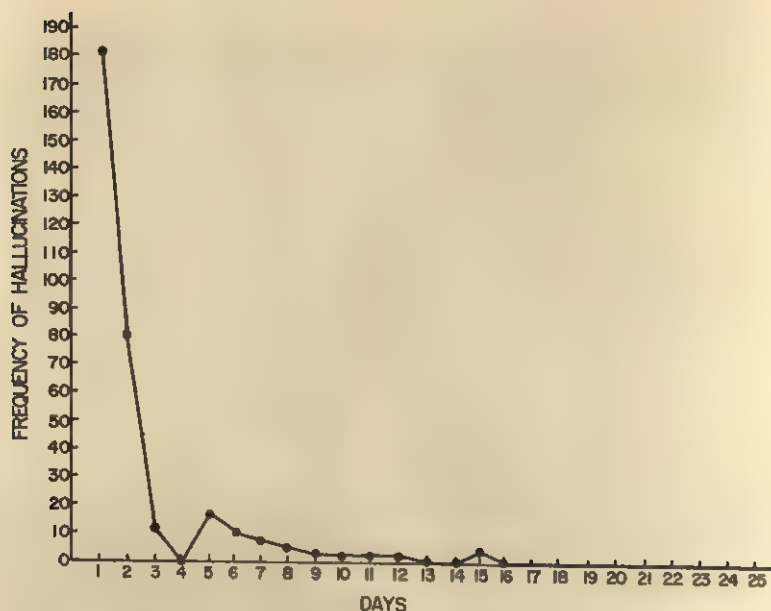


FIG. 1. Daily frequency of reported hallucinatory behavior.

fellow residents of the ward. The social reinforcement, consisting of praise, attention, encouragement, etc., was contingent on the recording of a reduction in reported hallucinatory behavior. When Eva failed to report a reduction, social reinforcement was withheld.

RESULTS AND DISCUSSION

Figure 1 represents the results of the present investigation. The chart identifies reported hallucinatory behavior on a daily basis. An inspection of Figure 1 reveals a marked reduction in this behavior. Reported hallucinations decreased in frequency from 181 on the first day, to 11 on the third day of private recording. No hallucinations were reported on the first day of public recording. Reported hallucinations then increased in frequency to 16 and decreased and leveled off at 0 on the sixteenth day. Reported hallucinatory behavior has remained at 0 with no remissions reported for approximately 6 months.

The above results raise a very provocative question: What could have accounted for this reduction in reported hallucinatory behavior? Since reported hallucinatory behavior steadily decreased in frequency during the 3 days of "private" recording, would "private" recording have been sufficient to maintain this reduction? While this question cannot be definitely

answered on the basis of the present study, it seems reasonable to assume that the chart placement and subsequent social reinforcement were responsible for maintaining the reduction in reported hallucinatory behavior. During the 3-day period prior to the recording chart being placed on the ward, Eva's hallucinations and their subsequent recordings were "private events" furnishing Eva with feedback exclusively. However, once the chart was provided, Eva's behavior became a "public event" subject to social approval and reinforcement. The remission reported on the fifth experimental day was most probably a reaction to this change in conditions. Concomitant with this reduction in reported hallucinations, a dramatic change was noted in other aspects of Eva's behavior. She appeared more relaxed, began to socialize more with other patients, her appearance improved, and she asked for a job. Perhaps more important, although certainly less quantifiable, she appeared to be happier. The results indicate that the present procedures were successful in reducing the frequency of reported hallucinatory behavior and in maintaining this reduction for an extended period of time. While these results are encouraging, they must be interpreted with some qualification. There is no evidence

to suggest that the results obtained with the present patient can be generalized to all patients exhibiting similar behavior. Self-monitoring, like other forms of self-control, is at present limited in its application. Since the self-control techniques require patient cooperation and participation in carrying out the procedures without direct supervision, it is necessary that the patient be capable of exhibiting the necessary behaviors to meet such requirements. Another variable is the degree of aversiveness of the maladaptive behavior to the patient. Unless this behavior is aversive to the extent that its reduction or termination is reinforcing to the patient, it is highly unlikely that patient participation can be maintained in the absence of the therapist.

The present report has been concerned with the frequency of reported hallucinatory behavior. It was this report of hallucinations that was observed and recorded as the overt maladaptive behavior. Since hallucinations are rarely if ever observable events, it might be appropriate to ask what, in addition to such reports, do clinicians mean when they say "the patient hallucinated?" If they mean only that the patient gave a verbal report of an hallucination, then it must be assumed that it is this verbal report that is considered to be the psychotic behavior.

It has been strongly suggested that in mental hospitals much of what is identified as psychotic behavior is nothing more than bizarre verbal behavior which is inadvertently

shaped up and maintained by the social reactions of the staff (Ayllon & Haughton, 1964). It would seem that if psychotic verbal behavior can so easily be shaped up and maintained by bootleg social reinforcement, then perhaps such behavior is more a reflection of the reinforcement contingencies found in the social environment than of some underlying intrapsychic conflict.

REFERENCES

- AYLLON, T., & HAUGHTON, E. Modification of symptomatic verbal behavior of mental patients. *Behaviour Research and Therapy*, 1964, 2, 87-97.
- GOLDIAMOND, I. Self-control procedures in personal behavior problems. *Psychological Reports*, 1965, 17, 851-868. (a)
- GOLDIAMOND, I. Stuttering and fluency as manipulatable operant response classes. In L. Krasner & L. Ullmann (Eds.), *Research in behavior modification*. New York: Holt, Rinehart & Winston, 1965. (b)
- KANTER, F. H. Self-regulation: Research, issues and speculations. Paper presented at the Ninth Annual Institute for Research in Clinical Psychology, University of Kansas, April 1967.
- KANTER, F. H., & PHILLIPS, J. S. Behavior therapy: A panacea for all ills or passing fancy? *Archives of General Psychiatry*, 1966, 15, 114-128.
- RUTNER, I. T. The modification of smoking behavior through techniques of self-control. Unpublished master's thesis, Wichita State University, 1967.
- STOLLACK, G. E. Weight loss obtained under various experimental procedures. Unpublished manuscript, Indiana University, 1966.
- SULZER, E. S. Research frontier: Reinforcement and the therapeutic contract. *Journal of Counseling Psychology*, 1962, 9, 271-276.

(Received for early publication March 5, 1969)

PERCEPTION OF DISABILITY BY THE NONDISABLED¹

A. P. MacDONALD, JR.,² AND JANET HALL

Rehabilitation Research and Training Center, West Virginia University

Fifty graduate students (30 males and 20 females) rated the seriousness of four disability types for six personal-social areas as they would affect an adult male, 28 years of age, married, with two preteenage children. Considerable support was obtained for the notion that certain disabling conditions are seen as more debilitating to some areas of life than others. Using Duncan's multiple-range test, 70% (i.e., 42 out of 60) of the pairs of personal-social dimensions were found to differ significantly. As predicted, locus of control and ratings of the seriousness of emotional disorders were significantly correlated ($r = -.30$). That is, emotional disorders were seen as more debilitating by internally controlled Ss. No support was found for the hypothesis that disabilities other than emotional disorders would be viewed as more debilitating by externally controlled persons. Implications of the findings for rehabilitation programs are discussed.

Numerous approaches are continually tried in an effort to rehabilitate greater numbers of clients. Notable among these is the approach directed at the discovery of especially relevant psychosocial variables. Once identified and understood, such variables may be instrumental to the development of more efficient intervention strategies, as well as to the refinement of existing strategies.

Though it is widely recognized and accepted that the disabled are disabled *people* rather than *disabled* people, remedial programs seem to focus more on the latter. Major emphases seem to be placed on (a) the offering of vocational rehabilitation services (which typically include vocational training and job placement) and (b) the treatment of the disabling condition itself (e.g., prosthetic devices and surgery). A broadening of focus may do much to facilitate the rehabilitation process.

A disabling condition may have, in addition to vocational impact, a measureable effect on a number of psychosocial dimensions. It seems proper to question the impact on such areas as family, social interaction, parenthood, marital

relationships, self-concept, etc., separately and in their various combinations. For example, if a specific disability has a greater effect on marital relationships than on vocation, effective vocational training may be inhibited by the use of in-residence training programs. In such cases, a more successful approach might be to keep clients within their families and to couple vocational training with marriage counseling.

It is also important to know how others view the disabled, as well as how the disabled view themselves. Both questions have been addressed. For example, the disabled and others have been found to perceive disabled persons as "different" from the nondisabled (Neff & Weiss, 1965). The perception of the disabled as different may have important implications for the ways in which society will treat them. One consequence is that they are often given social status comparable to that of an underprivileged minority group (Wright, 1960).

Apart from society's perception and treatment of the disabled, there is the issue of his perception of himself. Personality characteristics may in part determine the nature of that perception. Characteristics such as pessimism, low self-concept, etc., may be as crippling as any physical impairment, and may do much to inhibit adjustment to disability. For example, denial (i.e., refusing to accept the fact that one is disabled) is viewed by some investigators as the primary psy-

¹ This research was supported in part by West Virginia University Regional Rehabilitation Research and Training Center, through the Social and Rehabilitation Service, Department of Health, Education, and Welfare.

² Requests for reprints should be sent to A. P. MacDonald, Jr., Research and Training Center, Old Mountmainlair, Room 41, West Virginia University, Morgantown, West Virginia 26506.

chological mechanism which interferes with the rehabilitation process (Alger & Rusk, 1955; Ashenhost, Hurwitz, & Gruen, 1960; Barker & Wright, 1952; Barnes, 1952; Fisher, 1958). An important contribution would, therefore, be made if personality characteristics peculiar to denial of disability could be isolated.

One personality variable which would appear to be especially relevant to adjustment to disability is locus of control (Rotter, 1954, 1966). Locus of control deals with the extent to which persons view themselves as having some control over their reinforcements. Those who believe that they have control are considered to be internally controlled (I's). Those who believe that their destinies are controlled by luck, chance, fate, or powerful others, are externally controlled (E's). I's have been found to be (a) more willing to engage in remedial behaviors to confront their personality problems (Phares, Ritchie, & Davis, 1968), (b) more likely to stop smoking after the release of the research report that linked cancer to smoking (James, Woodruff, & Werner, 1965), and (c) superior to externals in actively seeking (Davis & Phares, 1967) and learning (Seeman, 1963) information relevant to problem solution.

Though internals have been found to be more open to remedial programs, they have been reported to be somewhat more defensive regarding negative professional interpretations of their personalities. That is, I's were found to recall less negative information about themselves than did E's (Phares et al., 1968). One explanation of this finding might be that I's see emotional disorders (assuming that negative personality interpretations suggest emotional disorders) as relatively more threatening to themselves. To the extent that emotional disorders imply a loss of inner control, they might be more threatening to I's, who believe that they have control to lose. On the other hand, E's might be more threatened by other disabilities, the stimulus characteristics of which they would view as eliciting negative reaction in the external world.

The following study was done to investigate the relevance of personal-social dimensions to disability groups as well as to test the notions

that (a) I's perceive emotional disorders as more debilitating than E's and (b) E's view disabilities other than emotional disorders as more debilitating than I's.

METHOD

Subjects

Data used in this study were collected from 50 (30 males and 20 females) graduate students at West Virginia University. All Ss had been administered the Rotter Internal-External locus of control scale (Rotter, 1966) 1 month earlier, as part of another study.

Procedure

A disability scale was developed for the measurement of Ss' perceptions of the seriousness of disability types for several personal and social dimensions. The scale was designed to tap five major classes of disability: (a) internal disorders—heart condition, back condition, and diabetes; (b) sensory disorders—blindness, deafness, and speech loss (muteness); (c) disfigurements—scarred face, being extremely overweight (obesity), and hunched back; (d) amputations—amputated leg, amputated arm, and amputated hand; (e) emotional disorders—having irrational fears, being extremely depressed, and being withdrawn.

The Ss rated each disability on a four-point (from "extremely debilitating" through "not much debilitating"). The score for each disability type was computed by summing the ratings of the three specific disabilities contained within it.

The questionnaire was preceded by the following instructions:

Very little is known about people's views toward disability, and we need your help in collecting information in this area. We would like *YOUR OPINION* about the seriousness of different disabilities. This is not a test. There are no right or wrong answers; we are interested in *your opinion* only. Below is a list of disabilities. For each disability, please mark an X in the category which best describes how serious each disability is. Do not omit any of the disabilities. Rate each one only once (mark only one X for each disability). Do not place any X's on the lines separating the categories.

In that disabilities affect different people in different ways, we will be asking you to indicate how serious each disability is on the job, in the marriage, in the community, and so on.

In answering, assume that the person disabled is: (1) the male head of a family; (2) 28 years old; (3) a high school graduate; (4) earning a modest income; and (5) the father of two pre-teenage children.

The Ss were asked to rate the seriousness of each disability separately for six personal and social

TABLE 1

DIFFERENCES BETWEEN THE MEANS OF EACH OF THE PERSONAL-SOCIAL DIMENSIONS
FOR INTERNAL DISORDERS

Dimension	Vocational	Marital	Social	Parental	Familial	Personal
Vocational		**	**	*	*	
Marital			**			*
Social				**	**	**
Parental						
Familial						
<i>M^a</i>	7.32	6.40	5.04	6.76	6.54	6.98

Note.—Data were analyzed by Duncan's multiple-range test for correlated means. $N = 50$.^a The higher the mean, the more debilitating.* $p \leq .05$.** $p \leq .01$.

dimensions (each dimension was rated on a separate page). The dimensions were as follows:

1. Vocational. "In your opinion, how seriously debilitating are the following for the adult male in his job?"

2. Marital. "In your opinion, how seriously debilitating are the following for the adult male in his relationship with his wife?"

3. Social. "In your opinion, how seriously debilitating are the following for the adult male in his relationship with his friends and neighbors?"

4. Parental. "In your opinion, how seriously debilitating are the following for the adult male in his role as a father?"

5. Familial. "In your opinion, how seriously debilitating are the following for the adult male in his relationship with his family in general?"

6. Personal. "In your opinion, how seriously debilitating are the following for the adult male in his feelings about himself?"

As aforementioned, it had been decided a priori that the 15 disabilities might be collapsed into five major classes or disability types. To test the validity

of the a priori clusters, intercorrelations between all the items (within the personal dimension only) were computed. The correlations were computed on data collected from 78 cases, of which the present sample was a part.²

Correlational analysis provided considerable support for the a priori decisions. Only two adjustments were needed: (a) blindness had to be dropped from the sensory disorder cluster, as it correlated too generally, and (b) disfigurements and amputations were collapsed into one category, which was called "cosmetic disorders." Subsequent to these adjustments, no item correlated more highly with an item of any cluster than it did with the items within its own cluster.

Reliability estimates (Cronbach's alpha coefficient) ($N = 50$) were computed for each disability cluster. The mean coefficients for each cluster averaged over the six dimensions were as follows: internal dis-

² The 78 Ss were taken from disparate settings. To insure homogeneity of sampling, it was deemed appropriate to use the data collected from one class of 50 graduate students for the major analyses.

TABLE 2

DIFFERENCES BETWEEN THE MEANS OF EACH OF THE PERSONAL-SOCIAL DIMENSIONS
FOR SENSORY DISORDERS

Dimension	Vocational	Marital	Social	Parental	Familial	Personal
Vocational		**	*			
Marital			**	**	**	**
Social				**	*	**
Parental						
Familial						
<i>M^a</i>	5.96	4.76	5.44	5.94	5.82	6.16

Note.—Data were analyzed by Duncan's multiple-range test for correlated means. $N = 50$.^a The higher the mean, the more debilitating.* $p \leq .05$.** $p \leq .01$.

TABLE 3
DIFFERENCES BETWEEN THE MEANS OF EACH OF THE PERSONAL-SOCIAL DIMENSIONS
FOR COSMETIC DISORDERS

Dimension	Vocational	Marital	Social	Parental	Familial	Personal
Vocational		*	**	**	**	**
Marital			*			**
Social				**	**	**
Parental						**
Familial						**
<i>M^a</i>	13.26	12.04	10.80	12.04	11.90	15.82

Note.—Data were analyzed by Duncan's multiple-range test for correlated means. $N = 50$.

^a The higher the mean, the more debilitating.

* $p < .05$.

** $p \leq .01$.

orders, .64; sensory disorders, .78; cosmetic disorders, .81; and emotional disorders, .85.⁴

RESULTS

Perception of Disability

A one-way analysis of variance was done for each disability type across the six personal-social dimensions separately. All F s were significant at well beyond the .005 level of confidence. To test for the differences between the means of each disability type across the dimensions, the data were analyzed with Dun-

⁴ The reliability estimates for each disability type for each of the dimensions (i.e., vocational, marital, social, parental, familial, and personal, respectively) were (a) internal disorders: .60, .69, .72, .62, .57, and .65; (b) sensory disorders: .65, .77, .83, .74, .83, and .84; (c) cosmetic disorders: .74, .82, .81, .84, .79, and .86; and (d) emotional disorders: .77, .80, .86, .84, .91, and .92.

can's multiple-range technique for correlated means.

The mean ratings obtained for each disability cluster, within each dimension, and the results of the Duncan multiple-range analyses (which tested the significance of the differences between the means for all pairs) are reported in Tables 1 through 4.

Inspection of the data reported in Tables 1 through 4 reveals considerable support for the notion that people alter their perceptions of disabilities according to specified dimensions. Seventy percent (i.e., 42) of the 60 pairs of dimensions tested were found to differ significantly.

The combinations that were found to differ significantly are too numerous to present in detail here—though they are, of course, summarized in Tables 1-4. Certain molar com-

TABLE 4
DIFFERENCES BETWEEN THE MEANS OF EACH OF THE PERSONAL-SOCIAL DIMENSIONS FOR
EMOTIONAL DISORDERS

Dimension	Vocational	Marital	Social	Parental	Familial	Personal
Vocational		**	**	**	**	**
Marital				**	**	
Social				**	**	
Parental					**	**
Familial						**
<i>M^a</i>	7.74	8.86	8.98	10.10	9.78	8.98

Note.—Data were analyzed by Duncan's multiple-range test for correlated means. $N = 50$.

^a The higher the mean, the more debilitating.

* $p < .05$.

** $p \leq .01$.

parisons may be briefly referred to, however. For convenience, they will be summarized by disability type:

Internal disorders: The vocational dimension received the highest mean rating. The mean for vocation differed significantly from the means given to all dimensions except the personal dimension. Internal disorders were seen to be least important socially. The mean rating for the social dimension was significantly lower than that of any other dimension.

Sensory disorders: The mean rating for the personal dimension was found to be highest. The personal dimension differed significantly from all dimensions except the vocational and parental. Seen as least affected was the marital dimension, which differed significantly from each of the other dimensions.

Cosmetic disorders: The mean rating assigned to the personal dimension was significantly greater than that given to any other dimension. Though the vocational area was rated as significantly less affected than the personal, it perhaps should be noted that it was rated as second highest and significantly higher than any of the remaining dimensions. The social dimension received a significantly lower rating than any of the other dimensions.

Emotional disorders: Contrary to the pattern reflected in the ratings assigned to the internal, sensory, and cosmetic disorders, the vocational dimension was rated as *least* affected by emotional disability. The mean rating assigned to the vocational dimension was significantly lower than those of all of the remaining dimensions. The parental dimension was seen as being significantly more affected by emotional disability than any of the other dimensions.

The investigators were somewhat concerned about the independence of the marital, parental, and familial dimensions. Since these dimensions all fall within the purview of the general rubric "family," it was conceivable that Ss would not distinguish between them. That is, though they were conceptually distinct, they may not have been empirically so. The data indicate that the Ss did distinguish between these dimensions. Within sensory disorders, the mean rating for the marital dimension differed significantly from those given

to the parental and familial dimensions. All three dimensions differed significantly from one another within the category of emotional disorders.

Locus of Control

Pearson product-moment correlations were done for the ratings of each disability type and the locus of control scale. To control for the effects of the six personal-social dimensions, as well as to minimize the number of statistical comparisons, the sum of the cluster scores over all dimensions was used. That is, the internal disorder score was the sum of the six ratings of internal disorders which each individual made. The same procedure was followed for the remaining disorders. Locus of control scores were available for 47 of the 50 Ss. As predicted, I's reported that they perceived emotional disorders as more debilitating than E's $A = -.30$ ($df = 45$, $p < .05$, two-tailed test) correlation obtained between the scores for emotional disorders and the locus of control scale (high scores are associated with externality). No significant relationship was found to obtain between locus of control and internal, sensory, or cosmetic disorders. The correlations were .16, -.05, and .01, respectively.

DISCUSSION

It remains for future research to determine how closely the attitudes of the nondisabled toward disabilities approximate those of the disabled. One interesting finding from these data is that the nondisabled generally perceived disabilities as being less debilitating socially than elsewhere. If these perceptions mirror those of the disabled, we may be over-emphasizing the social area at the expense of more crucial areas. If, on the other hand, these perceptions are highly discrepant from those of the disabled, these data are reflecting gross insensitivity to the problems experienced by the disabled. Validation of the former notion would indicate that emphasis might be best placed elsewhere. Validation of the latter would indicate a need for educating the nondisabled about the problems of the disabled. Such a program might do much to facilitate social adjustment among the disabled.

In this study, Ss were asked to report their perceptions of the debilitating effects of various disabilities on six personal-social dimensions. This is a departure from the usual research practice of having Ss report their feelings about the disabled. The present method (although it may have implications for determining attitudes toward the disabled) was chosen with specific interests in mind. Apart from the obvious concern for measuring attitudes of the nondisabled toward disabilities per se, the investigators were interested in gaining information about perceptions of the nondisabled regarding the effects of disabling conditions on important aspects of life. To the extent that such perceptions might determine one's reactions to becoming disabled, we may be in a position to draw valuable implications.⁵

Self-reported attitudes about disabled persons have been regarded as indexes of behavior toward the disabled. Similarly, self-reported attitudes about disabilities may be valid indexes of reactions to becoming disabled. It must be emphasized, however, that the validity of this assumption has not (to the authors' knowledge) been tested. Therefore, interpretation of these data, from that point of view, must be regarded as too highly speculative to be engaged in here. It remains for future research to clarify this issue.

The finding that Ss discriminated considerably between the personal-social dimensions seems to be an important one. Though it may turn out that the pattern of discrimination differs between disabled and nondisabled persons, it would seem that effective rehabilitation might be greatly facilitated if such discriminations were measured and considered. For example, it might be found that some disabled persons view their conditions as most debilitating to their marital relationships and least debilitating vocationally. In such cases the practice of bringing those clients into in-residence vocational training programs

might be ill advised. Rehabilitation of such clients would most likely be greatly facilitated by permitting them to remain with their families and by coupling vocational training with marriage counseling or family therapy. Correspondingly, a finding that a disability was seen to be most debilitating to a client in "his feelings about himself" would suggest that personal counseling or therapy might be an extremely important adjunct to that client's rehabilitation program.

At the very least, the findings here (though far from definitive) indicate a need for future research to explore the differential impact of disabilities on various personal and social aspects of life. Resulting knowledge may prove extremely valuable to the refinement and development of rehabilitation programs.

The hypothesized relationships between internal-external locus of control and reactions to the classes of disabilities received only mild support from these data. Though a significant correlation was found to obtain between locus of control and ratings of emotional disorders (as predicted, I's perceived emotional disorders as more debilitating than E's), none was found to be significant for the remaining disability types—though two of the three were in the predicted direction. Even the significant correlation obtained ($-.30$) was rather low, however. In that locus of control is an intrapersonal variable, the hypotheses might have been more efficiently tested had the Ss been asked to judge the debilitating effects of the disabilities in relation to themselves, rather than to another (i.e., the adult married male) (see Footnote 5).

REFERENCES

- ALGER, B., & RUSK, H. The rejection of help by some disabled people. *Archives of Physical and Medical Rehabilitation*, 1955, 36, 279-281.
- ASHENHORST, E., HURWITZ, L., & GRUEN, A. Motivational and structural factors in the denial of hemiplegia. *Archives of Neurology*, 1960, 3, 315-317.
- BARKER, R. G., & WRIGHT, B. The social psychology of adjustment to physical disability. In J. F. Garrett (Ed.), *Psychological aspects of physical disability*. Washington, D. C.: Department of Health, Education, and Welfare, Office of Vocational Rehabilitation, 1952.
- BARNES, R. H. Psychological problems in physical rehabilitation: A review. *American Journal of Medical Sciences*, 1952, 223, 106-112.

⁵ It should be kept in mind that the Ss were asked to indicate how debilitating various disabilities would be for another person. Caution should therefore be exercised in generalizing from responses about others to responses about self. The authors are presently conducting a study in which Ss are asked to rate the seriousness of the disabilities in relation to self.

- DAVIS, W. L., & PHARES, E. J. Internal-external control as a determinant of information-seeking in a social influence situation. *Journal of Personality*, 1967, 35, 547-561.
- FISHER, S. Mechanism of denial in physical disabilities. *Archives of Neurology and Psychiatry*, 1958, 80, 784.
- JAMES, W. H., WOODRUFF, A. B., & WERNER, W. Effect of internal and external control upon changes in smoking behavior. *Journal of Consulting Psychology*, 1965, 29, 127-129.
- NEFF, W. S., & WEISS, S. A. The psychological aspects of disability. In B. A. Wolman (Ed.), *Handbook of clinical psychology*. New York: McGraw-Hill, 1965.
- PHARES, E. J., RITCHIE, D. E., & DAVIS, W. L. Internal-external control and reaction to threat. *Journal of Personality and Social Psychology*, 1968, 10, 402-405.
- ROTTER, J. B. *Social learning and clinical psychology*. New York: Prentice-Hall, 1954.
- ROTTER, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 1966, 80(1, Whole No. 609).
- SEEMAN, M. Alienation and social learning in a reformatory. *American Journal of Sociology*, 1963, 69, 270-284.
- WRIGHT, B. A. *Physical disability—A psychological approach*. New York: Harper & Row, 1960.

(Received for early publication May 19, 1969)

FREQUENCY OF DREAM RECALL ESTIMATED BY THREE METHODS AND RELATED TO DEFENSE PREFERENCE AND ANXIETY¹

DAVID BENJAMIN COHEN²

University of Michigan

Dream recall frequency (DRF) was measured by questionnaire (Q), diary (D), and sleep interruption via telephone (T). Q DRF was high for aggression out, projection, and denial, and low for intellectualization and aggression in ($p > .05$), especially for Ss higher on test anxiety ($p < .05$) than the median of their particular defense group (high "relative anxiety"). Compared to Q DRF, D and T DRF were markedly higher for all groups; there were no defense differences. The Ss with higher relative anxiety had significantly more D and T DRF than Ss with low relative anxiety ($p < .05$). There were no Q, D, or T DRF differences for high versus low "general anxiety" (anxiety relative to the overall sample median). The relationship of DRF to procedural, psychological, and psychophysiological factors is discussed.

Until quite recently, the focus of clinical and experimental investigation of dreaming has been on (a) interpretation and standardization of dream content (Freud, 1955; Hall, 1966); (b) psychophysiological concomitants of "dreaming sleep" (Hartmann, 1967; Luce, 1966; Snyder, 1963); (c) psychological aspects of sleep and dreaming (Foulkes, 1966); and (d) hypothesizing about the possible biological functions of various stages of sleep (Fisher, 1965; Shapiro, 1967; Ullman, 1959).

In light of the large amount of empirical work on the psychophysiology of sleep and dreaming (Hartmann, 1967), there has been relatively little systematic investigation of psychological factors contributing to differential frequency of dream recall (DRF). In the absence of a definitive psychological theory of DRF, the few studies reported in the literature have focused on variables that seem logically related to DRF. A small but significant relationship has been found between high test anxiety and high DRF measured by questionnaire (Domhoff & Gerson, 1967;

Strauch, Bliemeister, & Malzahn, 1966; Tart, 1962), and high DRF measured by home dream diaries (Singer & Schonbar, 1961). Schonbar (1959) reported a similar relationship between test anxiety and diary DRF, but the unusually high correlation of .50 may have been due to statistical artifact (Domhoff & Gerson, 1967).

Negative correlations between DRF and ego strength were found by Schonbar (1959) using diaries, and by Domhoff and Gerson (1967) using questionnaires. Tart (1962) and Domhoff and Gerson (1967) found low but significant negative correlations between questionnaire DRF and repression. In light of the nearly tautological nature of the prediction, the small proportion of the variance accounted for by repression is surprising and challenging.

Other studies have reported low but significant correlations between questionnaire and diary DRF and cognitive control measured by the Embedded Figures Test (Witkin, Dyk, Faterson, Goodenough, & Karp, 1962), the rod-and-frame test (Schonbar, 1965), and the schematizing test (Lachman, Lapkin, & Handelman, 1962). Measures of "fantasy predominance," "associative productivity," and "introspective constriction" derived from Rorschach scores (Orlinsky, 1966) were correlated with DRF measured by the rapid eye movement, sleep interruption method of Dement and Kleitman (1957). The first two

¹ This paper is based on a dissertation submitted to the University of Michigan in partial fulfillment of the requirements for the PhD degree. The author is especially grateful to Arthur Shapiro of the Downstate Medical Center, New York City, and Harry Gollob of the University of Michigan for their suggestions and support.

² Requests for reprints should be sent to David B. Cohen, who is now at the University of Texas, Department of Psychology, Austin, Texas 78712.

Rorschach measures correlated positively, the last one negatively with this measure of DRF.

Empirical research has suggested a variety of psychological, psychophysiological, and methodological factors which seem to contribute to DRF: the method used (e.g., checklist vs. verbatim diary), the utility of recalling dreams, consolidation of memory traces, attitude toward dream life, repression, the psychophysiology of the waking process, receptivity toward inner life, mislabeling dreams as thoughts, "depth" of sleep, stage of sleep, anxiety, confabulation of "recall" upon awakening. Therefore, in the study reported here, DRF is treated as a dependent variable.

Relationship of DRF to the Method

The questionnaire (Q) method is derived from a set of global estimations of a collection of past events, that is, estimation by *S* of a preconscious, experiment-independent tendency to recall dreams upon waking from sleep. Diary (D) and telephone (T) measures of DRF are frequency counts of day to day, experiment-dependent (i.e., externally imposed) and therefore conscious focusing of attention on the psychological aspects of the waking process. It was assumed that for most *Ss* the absolute amount of DRF would increase from initial Q estimates to D and T measures. Complex shifts, rather than monotonic increases, in the relative amount of DRF from Q to D and T as a function of personality differences would underscore the importance of specifying method when discussing the relation of DRF to personality.

Relationship of DRF to Defense Preference and Anxiety

Psychoanalytic theory acknowledges the close relationship of defense and anxiety. Defense is conceptualized as an adjustive means of containing ("countercathecting"), binding, and/or "neutralizing" (i.e., sublimating) anxiety. Defense styles could be thought of as the psychologically "bound" form of free anxiety in the sense that matter is a bound form of energy. There is strong theoretical justification for focusing on the relationship or *interaction* of "energetic" (free anxiety) and "structural" (defense style) aspects of normal

and neurotic adjustment rather than looking at these factors separately. There is generally a similar approach to behavior analysis in the Hullian concept of the interaction of drive and habit strength.

In most studies, DRF has been related, not only to isolated dimensions, but relatively undifferentiated "energetic" and "structural" categories (e.g., anxiety, neuroticism, etc.) that tend to obscure potentially relevant individual differences. In the study reported here, the "structural" dimension, defensiveness, was operationalized with a defense preference test by which five groups of *Ss*, each characterized by a different defense style, can be obtained. The five defense groups are aggression toward objects, projection, intellectualization, aggression toward the self, and denial-repression.

Anxiety may be statistically independent of "defensiveness" or other general structural personality variables, but it is dependent, in both a statistical and a psychological sense, on *particular* defensive styles. Each style has an expected or "normal" level of anxiety, and it is this level, rather than some overall population value, which provides a meaningful criterion against which high or low levels of anxiety are predictive of behavior. For example, the median value of anxiety test scores of groups high on "intellectualization" might be in the vicinity of 20, while a broadly defined, structurally less differentiated group high on "defensiveness" might obtain a median value around 30. If nothing else were known about an *S* other than that he is high on "defensiveness," his anxiety test score of 30 would have little predictive value. If it were known that this *S* could be clearly classified as an "intellectualizer," that same score of 30 might indicate high anxiety relative to what his particular defensive style can comfortably accommodate. The score of 30 might be diagnostically indicative of a serious weakening of defensive capacity.

Therefore, in the present study, all *Ss* with anxiety test scores above the median of their particular defense group were considered to be high "relative anxiety" *Ss* whatever the relationship of their score to the overall sample median, that is, whatever their level of "general anxiety." The *Ss* with anxiety test scores below the median of their particular defensive

group were called low "relative anxiety" Ss.

Accordingly, the major prediction was that DRF would be most clearly related to level of relative anxiety, and only minimally if at all related to general anxiety. No specific predictions were made as to the differences in the relationship under the three methods of obtaining DRF.

METHOD

Subjects

The Ss were 86 male undergraduate students most of whom were freshmen and sophomores, all but a few of whom lived in dormitories. They were selected from an S pool made up of students taking introductory psychology courses at the University of Michigan, and thus were neither paid nor volunteers.

Materials

Independent variables. The Defense Measuring Instrument (DMI) devised by Gleser and Ihlevich (1969) was used to select Ss for five defense preference (experimental) groups and one defense non-preferent (control) group. The five defense dimensions scored are turning aggression outward toward others (TAO), projection (PRO), intellectualization (INT), turning aggression inwardly onto the self (TAS), and reversal (REV, i.e., denial). According to the authors of the test, these dimensions represent a comprehensive and theoretically meaningful way of dividing the field of defense styles. The test is an objectively scored, highly structured projective test involving responses to stories tapping various conflict areas. A number of construct validity measures have been obtained (Gleser & Ihlevich, 1969).

The IPAT Anxiety Test developed by Cattell and his associates was used to obtain anxiety scores. This is an objective, 40-item questionnaire with very high reliability and "varied and impressive" evidence of validity (Buros, 1965).

Dependent variables. Three measures of DRF were used: (a) estimation of frequency of recall from questionnaire items chosen by S (Q DRF); (b) frequency count of indications of dream recall on a 30-day check list "diary" sheet (D DRF); and (c) frequency count of indications of dream recall for seven completed calls during the period of about 5 weeks (T DRF). Any perceptual, conceptual, or affective element presumed by S to be part of a dream was scored "dream recalled."

A discussion of the relationship of dream experiences, dream recall, and dream reporting is beyond the scope of this report.³ It is assumed, however, that reports of dreaming more often than not are genuine indications of the presence of recall, and

that recall is generally a residue of an actual dream experience rather than some sort of arousal phenomenon (Foulkes, 1967) or confabulation to please E (Rechtschaffen, 1967).

The questionnaire (Q) was composed of nine items, but only the item on frequency of occurrence was used. The S was asked to choose one of four categories of frequency as most characteristic of himself: (a) just about every day or every other day; (b) about once a week or twice a week; (c) couple of times a month; (d) hardly ever. The S could pick a second-most-applicable item. Verbal estimations were converted into numerical estimations based on a month's time. In a separate study of 23 Ss taken from a similar student population, a test-retest product-moment correlation of .906 was obtained for an interval of 4 weeks.⁴

The dream "diary" (D) was an objective checklist composed of six of the nine questions from Q plus a question on length of time between waking and filling in the sheet. The S responded by putting a mark (X) under the appropriate "a versus b" alternative for each question within the column that represented that particular day. The sheet was divided into 30 columns, one for each of the 30 days S was to fill out the sheet. As with Q, only the item contributing to DRF was used.

A split-half product-moment correlation of .92 obtained for the 23 Ss of the control group indicated high stability over time (i.e., from the first to the second 15-day interval). The odd-even correlation for the same was .65 ($p < .001$) indicating a certain amount of day to day fluctuation. There seems to be a general consistency of DRF under conditions that impose an ongoing task of reporting the psychological aspects of waking.

The telephone procedure (T) to be described was based on the same six questions from Q that were used in D. A "telephone diary sheet" was filled out by the research assistants (RAs) who made the calls.⁵ While scoring the interview on the T sheet, RA recorded the interview with a Wollensak T-1500 tape recorder by means of a pickup coil placed near the earpiece of the receiver. The transcriptions of these interviews were not used in this study. Only the item relevant to DRF was used. The T procedure was suggested by Arthur Shapiro,⁶ and used in a study by Rosen (1967).

Procedure

The Ss were selected for the five defense groups if their score on a particular dimension was at least equal to or greater than the mean plus 1.5 standard deviations, and was the highest score of the five. The Ss were selected for the control group by randomly choosing 23 Ss, all of whose scores were within the expected range of ± 1.25 standard devia-

⁴The author wishes to thank Gary Talpos for providing the data for the statistical test.

⁵Special thanks to Kurt Hecox and Andy Broder for their valuable assistance.

⁶A. Shapiro, personal communication, March 1967.

³A discussion of this relationship is planned in the context of another paper on the validity of the telephone (T) method of obtaining DRF data.

TABLE 1
INTERCORRELATIONS OF DRF FOR
THE THREE PROCEDURES

Procedure	Q	D	T
Q			
D		.01 ^a	.20 ^b
			.54 ^{b*}

^a $N = 83$.

^b $N = 81$.

* $p < .001$.

tions. After selection, Ss filled out the IPAT test and the dream questionnaire, and then received the dairy sheets.

Finally the T method was outlined. The Ss were told that during the period they were filling out the diary, they would be called sometime during the night once or twice a week, but not on week-ends. The Ss were not informed that the calls would be made only between the hours of 5:00 and 7:30 A.M., during which time Ss spent most of their time in dreaming sleep (Stages 2 and 1-REM sleep). The Ss were told to consider the D and T procedures independent, and not to skip D on the days they received a call (unless, of course, they couldn't go back to sleep).

An extended discussion of the relationship of T to laboratory sleep interruption methods is being prepared for publication. Suffice it to say there is considerable evidence that in terms of both DRF and content of dream recall, there is great similarity in the data obtained by home and laboratory methods. If one is interested in DRF or content of dream recall rather than the relationship of these to the particular stage of sleep from which S is aroused, T fairly well simulates the laboratory procedure.

RESULTS AND DISCUSSION

Relationship of DRF to Method

In order to explore the meaning of DRF as variously measured, intercorrelations and analysis of variance of DRF from the three methods were obtained. Table 1 shows the intercorrelations of DRF for questionnaire (Q), diary (D), and telephone (T) procedures. Only the correlation between D and T is significant. The variance shared by these measures is probably due to the fact that they both involve on the spot, continuous, externally imposed, conscious focusing of attention on an otherwise preconscious process. These intercorrelations suggest that Q represents a different set of factors from D and T, and that although D and T are far from identical, they do seem similar.

Table 2 bears out this initial impression. Table 2 shows mean DRF for the three methods when DRF is observed as a function of defense preference and anxiety. Overall estimations of DRF increased from about six times per month for Q to about 18 times per month (about 60%) for D, and to approximately one out of every two calls (about 55%) for T. Note the similarity of percentage for D and T.

From the analyses of variance carried out on the data of Table 2 (see Table 3), it is fairly clear that the D and T data show a greater similarity to each other than either does to the Q data, both in terms of which effects are significant and the proportion of variance accounted for. Specifically, the defense (just short of the .05 level) and interaction ($p < .05$) effects are significant and account for about 19% of the variance in the Q DRF data. Ordinarily, an effect failing to reach the .05 level would be most parsimoniously reported as nonsignificant. However, in light of the apparently close relationship between anxiety and defense, both in terms of

TABLE 2
MEAN DRF RELATED TO DEFENSE PREFERENCE AND
ANXIETY FOR Q, D, AND T METHODS

Level of anxiety	Defense preference groups					Control
	TAO	PRO	REV	INT	TAS	
Questionnaire						
High	8.2	8.9	7.7	1.6	2.8	6.8
Low	5.8	3.9	5.3	3.2	4.3	5.4
Diary						
High	21.7	23.0	23.1	14.2	18.9	20.2
Low	13.6	14.1	15.5	16.6	18.8	14.9
Telephone ^a						
High	14.6	15.0	19.3	13.7	18.4	18.4
Low	10.7	12.9	9.4	13.7	11.6	14.6

Note.—Cell Ns for the defense preference groups vary from five to seven; for the control group, 10 to 11.
^a Cell numbers are mean DRF per 7 calls prorated to represent DRF per 30 calls. This allows direct comparison of numbers in the three tables.

TABLE 3

ANALYSES OF VARIANCE OF DRF RELATED TO
DEFENSE PREFERENCE AND ANXIETY FOR
Q, D, AND T METHODS

Source	df	MS	F	η^2
Questionnaire				
Relative anxiety	1	36.5	1.9	
Defense preference	4	46.1	2.4*	.09
Defense \times Anxiety	4	49.9	2.6**	.10
Error (within groups)	52	19.2		
Diary				
Relative anxiety	1	362.6	4.4**	.07
Defense preference	4	15.5	—	
Defense \times Anxiety	4	104.1	1.3	
Error (within groups)	48	83.0		
Telephone				
Relative anxiety	1	21.6	4.6**	.07
Defense preference	4	0.8	—	
Defense \times Anxiety	4	2.5	—	
Error (within groups)	46	4.7		

Note.—Cell Ns for the defense preference groups vary from five to seven; for the control group from 10 to 11. η^2 is an unbiased estimate of the proportion of variance accounted for (Cohen, 1965).

* $p > .05$.

** $p < .05$.

theory and data reported in this study, it is felt that this defense effect is significant.

Neither D nor T data show such variability. For these methods, only the relative anxiety effect is significant ($p < .05$), accounting for about 7% of the variance in each set of data. Apparently the demand characteristic common to both procedures, that is, externally imposed focusing on dream life, can account for much of the D and T results. It is clear, though, that DRF cannot be considered independently of the measure used to obtain it, and that differences especially between Q versus D and T cannot be ignored, especially in studies which select Ss on the basis of high or low DRF.

Q DRF seems to be a function of defense-related, perhaps attitudinal, experiment-independent tendencies to focus or not to focus on dream life. D and T DRF seem to be relatively more independent of specific defense styles. They seem more related to an

experimentally induced, conscious tendency to recall dreaming. For all three methods, these "internally" or "externally" based tendencies are apparently exaggerated by high relative anxiety.

Relationship of Q DRF to Defense Preference

Defense Groups TAO (aggressive outward), PRO (projection), and REV (denial) showed significantly more Q DRF than Defense Groups INT (intellectualization) and TAS (turning aggression inwardly). When these groups were divided on the basis of high versus low relative anxiety, the relationship between defense styles and Q DRF increased; relatively anxious Ss in the aggression-out (TAO), projection (PRO), and denial (REV) groups had significantly more Q DRF (about eight times per month) than relatively anxious Ss in the intellectualizer (INT) and aggression-in (TAS) groups (about two per month). There was almost no differentiation in Q DRF across defense groups for relatively nonanxious Ss. Apparently, this lack of differentiation for low relative anxiety Ss contributed to the failure of the overall defense effect to reach the .05 level.

The importance of valuing dreams was suggested by Domhoff and Gerson (1967). Evidence for the relationship of Q DRF to a positive valuation of dream recall was indicated in results from the Strauch et al. study (1966) showing that Q recallers more often believe that dreams have meaning and like to recall dreams. Q DRF might be a reflection of the strength of S's identification with specific class-bound or subcultural attitudes toward dream life. Intellectualization (INT) and aggression-in (TAS) could be seen as a tendency toward overcommitment (in the sense of conformity, bureaucratization, etc.) to social undervaluation of the utility or relevance of dream life.

Perhaps dream recall is a frequently occurring, natural phenomenon of early childhood which is differentially reinforced by the socialization practices in societies that tend to value or undervalue dream recall. In our culture, does the presence of high Q DRF indicate a certain "primitivity" which may or may not be necessarily neurotic? In other

words, does high recall in some people indicate a certain freedom (e.g., nonconformity, creativity, capacity for innovation) from bureaucratic and "institutional" inhibition while indicating in others a neurotic rebelliousness, incompetence, or other disturbance?

Relationship of DRF to Relative Versus General Anxiety

It was predicted that DRF would be differentially related to *relative* anxiety rather than *general* anxiety. This prediction was borne out by the following results: (a) the relative anxiety main effects in the D and T data were significant ($p < .05$); (b) the interaction effect between relative anxiety and defense preference in the Q data was significant ($p < .05$); and (c) there were no significant differences in Q, D, or T DRF for high versus low *general* anxiety, either for the defense-preferent Ss as a whole (all of the Ss in the defense groups) or for the defense-non-preferent (control) Ss; none of the six t tests reached significance.

The differential effect of relative anxiety was also expressed in hostility elicited by the T procedure. Hostility was defined operationally on the basis of RAs' notes to the effect that S had hung up the phone without answering; expressed hostility verbally via sarcasm, impatience, exasperation; or was "not there" or "awake" on two or more separate occasions. A frequency count was made of Ss expressing hostility at least once during T. There were no defense differences. When the sample was divided on the basis of *general* anxiety, a chi-square of .29 ($df = 1$) was obtained, indicating no difference in the number of Ss high or low in general anxiety expressing hostility toward T. When the medians for individual defense groups were used to divide the sample, that is, when relative anxiety was used, a chi-square of 10 ($df = 1$) was obtained indicating a significant difference beyond the .005 level. Fewer Ss high in *relative* anxiety expressed hostility than Ss in the low relative anxiety group.

REFERENCES

- BUROS, O. K. (Ed.) *The sixth mental measurement yearbook*. Highland Park, N. J.: Gryphon, 1965.
- COHEN, J. Some statistical issues in psychological research. In B. B. Wolman (Ed.), *Handbook of clinical psychology*. New York: McGraw-Hill, 1965.
- DEMENT, W. C., & KLEITMAN, N. The relation of eye movements during sleep to dream activity: An objective method for the study of dreaming. *Journal of Experimental Psychology*, 1957, 53, 339-346.
- DOMHOFF, W., & GERSON, A. Replication and critique of three studies on personality correlates of dream recall. *Journal of Consulting Psychology*, 1967, 31, 431.
- FISHER, C. Psychoanalytic implications of recent research on sleep and dreaming. *Journal of the American Psychoanalytic Association*, 1965, 13, 197-303.
- FOULKES, D. Nonrapid eye movement mentation. *Experimental Neurology*, 1967, Supplement 4, 28-38.
- FOULKES, D. *The psychology of sleep*. New York: Scribner's, 1966.
- FREUD, S. *The interpretation of dreams*. New York: Basic Books, 1955.
- GLESER, G. C., & IMILEVICH, D. An objective instrument for measuring defense mechanisms. *Journal of Consulting and Clinical Psychology*, 1969, 33, 51-60.
- HALL, C. S. *The meaning of dreams*. New York: McGraw-Hill, 1966.
- HARTMAN, E. *The biology of dreaming*. Springfield, Ill.: Charles C Thomas, 1967.
- LACHMANN, F. M., LAPKIN, B., & HANDLEMAN, N. S. Recall of dreams, its relation to repression and cognitive control. *Journal of Abnormal and Social Psychology*, 1962, 64, 160-162.
- LUCE, G. G. *Current research on sleep and dreams*. Bethesda: Department of Health, Education, and Welfare, 1966.
- ORLINSKY, D. E. Rorschach test correlates of dreaming and dream recall. *Journal of Projective Technique and Personality Assessment*, 1966, 30, 250-253.
- RECHTSCHAFFEN, A. Dream reports and dream experiences. *Experimental Neurology*, 1967, Supplement 4, 15-28.
- ROSEN, G. Dream reporting in the home environment. Unpublished master's thesis, Brooklyn College, 1967.
- SCHONBAR, R. A. Some manifest characteristics of recallers and nonrecallers of dreams. *Journal of Consulting Psychology*, 1959, 23, 414-418.
- SCHOBAR, R. A. Differential dream recall frequency as a component of life style. *Journal of Consulting Psychology*, 1965, 29, 468-474.
- SHAPIRO, A. Dreaming and the physiology of sleep. A critical review of some empirical data and a proposal for a theoretical model of sleep and dreaming. *Experimental Neurology*, 1967, Supplement 4, 56-81.

- SINGER, J. L., & SCHONBAR, R. A. Correlates of dreaming: A dimension of self awareness. *Journal of Consulting Psychology*, 1961, 25, 1-6.
- SNYDER, F. The new biology of dreaming. *Archives of General Psychiatry*, 1963, 8, 381-391.
- STRAUCH, I., BLIEMEISTER, B., & MALZAHN, P. Spontaneous dream recall and some personality characteristics. Paper presented at the meeting of the Association for the Psychophysiological Study of Sleep, Gainesville, Florida, March 1966.
- TART, C. T. Frequency of dream recall and some personality measures. *Journal of Consulting Psychology*, 1962, 26, 467-470.
- ULLMAN, M. The adaptive significance of the dream. *Journal of Nervous and Mental Disease*, 1959, 129, 144-149.
- WITKIN, H. A., DYK, R. B., FATERSON, H. F., GOODENOUGH, D. R., & KARP, S. A. *Psychological differentiation*. New York: Wiley, 1962.

(Received July 10, 1968)

INTELLIGENCE DEFINED AS EFFECTIVENESS OF APPROACHES

CONSTANCE T. FISCHER¹

Duquesne University

The need for a generally acceptable definition of intelligence is reviewed. Emphasis is placed on the fact that any proposed definition should be consonant with research findings as well as current and previous conceptions. The following definition is offered: Intelligence refers to the effectiveness, relative to age peers, of the individual's approaches to situations in which competence is highly regarded by the culture. This proposed definition opens the way to a theoretical and clinically applicable conception of intelligence, which is also presented. Its main tenet is that along with more widely acknowledged variables, a person's available frames of reference or meaning networks are the limiting conditions of his possibilities for action. Within this scheme, frames of reference together with their behavioral aspect constitute the "approach" referred to in the definition.

The term "intelligence," as a referent to measured intellect, came into being in the wake of last century's faculty psychology. Subsequent formulations of intelligence and measurements thereof have been explicitly described by Tuddenham (1963). The guiding concerns were primarily based on the empirical, pragmatic activities of Binet and Terman. Thus, such issues as IQ constancy and culture-fair testing are by-products of the prevailing emphasis on test scores. The nature of intelligent behavior itself has received significantly less attention. Similarly, discussions of the nature-nurture controversy have centered on changes in test IQ rather than on the child and his environment. Factor analytic conceptions of intelligence, still the most prevalent, are, of course, restricted by the limits of already constructed, usually pragmatic tests. Perhaps the most poignant reminder of the difficulty in moving beyond the tests themselves is the continuing prominence of Boring's (1923) definition of intelligence as that which intelligence tests test.

The above operational *definition* has, of course, proved conducive to highly developed factor analytic *conceptions* of intelligence both as a hypothetical construct and as an intervening variable. During the development, definition and conception have become nearly

synonymous. *Definition* is intended here as the identification of a phenomenon as it is most widely understood. *Facts* are highly reliable observations of relevant relationships and may be placed in appendixes to the definition. The facts are about the defined phenomenon, and, of course, should be taken into account by a theory of *conception*. The latter, which goes beyond the facts, serves as a guide to integration and further fact-gathering. A definition can be consonant with many conceptions; a concept, however, necessarily precludes assimilation of some facts. A major tenet of the present paper is that while definition and conception imply each other, treating them as identical is not conducive to the development of theory and the discovery of empirical relationships.

Although there is, then, a need for an agreed-on definition of intelligence prior to research and theorizing, there is none which is commonly held and utilitarian. Theoretical articles usually proceed without an explicit definition of their topic. Psychology and education textbooks typically discuss the problems of definition, make reference to common-sense notions of brightness, offer the operational alternative under the heading of intelligence as a hypothetical construct, and then present a sample of theories from Binet through the latest factor-analytic conclusions. Finally, a summary of research facts is given. Where a straightforward definition of intel-

¹ Requests for reprints should be sent to Constance T. Fischer, Psychology Department, Duquesne University, Pittsburgh, Pennsylvania 15219.

ligence is attempted, the usual outcome is either a list of traits, said to represent intelligence, or an outright conception or theory of intelligence. In all cases, an explicit, commonly held definition is somehow by-passed. Where the operational alternative is given as such, it is not usually accompanied by what would seem to be a necessary discussion of intelligence tests.

The purpose of the present paper, then, is to offer a definition of intelligence that will reflect its most basic common meaning. At the same time, it will neither include nor exclude specific theories or conceptions. The proposed definition, nevertheless, is intended to reflect recent trends in both research and theory, and to encourage further developments. It is especially structured to allow for growing clinical conceptions of intelligence (e.g., Glasser & Zimmerman, 1967; Lyons, 1963), one of which is outlined.

PROPOSED DEFINITION

Intelligence refers to the effectiveness, relative to age peers, of the individual's approaches to situations in which competence is highly regarded by the culture.

This definition is consistent with the history of the concept as comparative, quantitative, and operationally grounded in standardized tests. It is thus consistent with the prevalent philosophy of science position, as presented by Spiker and McCandless (1954). They argue that intelligence is a useful concept insofar as it relates to other psychological variables, and that it has no reality outside of these relationships. Confusion about the nature of intelligence, according to their view, arises through the multiplicity of measurement operations. The proposed definition goes beyond abstract operational reference by specifying what it is that intelligence tests test, as well as by implying that measurement is not of an entity.

Besides serving historical continuity, the definition is contemporary. It is consistent with Bouchard's (1968) description of an emerging research emphasis on both the process character of intelligence and on the question of intelligence as "a sub-problem associated with the more general problem of determining the structure of cognition and

thought [p. 27]." Thus, since "structure of cognition and thought" may be subsumed under "approach," such research topics as information-processing strategies (Hunt, 1961), computer simulation of cognitive structures (Reitman, 1965), and a structure-of-intellect model (Guilford, 1967) may all be classified under intelligence as defined here. Piaget's (1950) formulations of schema, concept, assimilation, and adaptation may also be encompassed by "approach."

Effectiveness

The reference to effectiveness is consonant with both historical and current use of the term "intelligence"; that is, it brings to mind the early emphasis on problem solving, and is quite in line with recent proposals of intelligence as sets of abilities and/or adaptations (Hayes, 1962; Mischel, 1968; Tuddenham, 1963; Wallace, 1966; Wesman, 1968). From varying perspectives, all these authors stress the roles of antecedent conditions and reward parameters in the acquisition and demonstration of abilities (intelligence). The present definition's reference to effectiveness undercuts entity and constancy notions, which these theorists decry. It encourages reference to an individual's effectiveness rather than to his potential or aptitude. Indeed, this terminology implies what in practice is meant by aptitude: that present levels of effectiveness are generally predictive of later levels insofar as extant equipment, maturation, and experience are requisite to further development.

The reference to age-peer comparison is a reminder that in practice, IQ is the estimate of intelligence, and that it has specific statistical properties. Finally, reference to age-peer standing also suggests that type of effectiveness may change with age.

Approaches to Situations

In order for the definition to have maximal utility, the term "approaches" must be amenable to several interpretations. Specifically, strict behaviorists will conceive of approach as response. However, while the definition is intended to be useful to the behaviorist and learning theorist, its major implications are for the psychologist with an existentially oriented, holistic perspective. For the latter,

approach connotes both the frames of reference, or network of meanings, and intentional behaviors of the testee. Portions of this section, therefore, will not be relevant in their present form to behavioristic conceptions of intelligence.

Use of the term "approach" implies that there may be alternative routes to effectiveness. Inasmuch as the present definition avoids any implication that intelligence is an entity, the possibility of change is more readily apparent. Furthermore, the definition directs attention to the specific approaches used by the testee. In turn, it becomes possible to investigate ways of teaching him more effective approaches.

This view is also consonant with the empirical finding that test scores are unstable while patterns of perception and behavior are developing; concretely, up to and through the early school years, approaches are themselves unstable. Liverant (1960), in his presentation of intelligence as a descriptive rather than explanatory concept, emphasizes the need "to examine first the nature of antecedent conditions and then the subsequent responses subsumed by intelligence [p. 102]." Such an examination, he points out, would indicate whether a testee's failure of an item were due to expectancy and reward-value variables or to the absence of the correct response in the testee's repertory. In the present paper, however, it is posited that description must go beyond operationally identified stimulus and response variables. In this view of approach, the testee behaves in accordance with his personal frames of reference—the meanings of the stimulus object available and salient to him. It is, then, more fruitful to determine the identity of the stimulus (its meaning) for the testee. Similarly, it would be more productive to ask, first, what are the essential components (including frames of reference) of the ability to earn credit for a given item? And then, which components are absent for this testee?

The testee's situation includes more than a physical setting and a relationship to the examiner. The meaning, for him, of being tested is also a highly relevant frame of reference. For example, when a person fails an item it does not necessarily indicate

"inappropriate expectancies" or "reward values," or absence of the scorable response from his repertory. It may be the case that the testee has approached the test situation with a frame of reference quite different from that intended in the test manual. For instance, he may have viewed intelligence as creativity, and hence disdained to give obvious answers. Or, the presence of the examiner may have served the testee's prevailing concern with interpersonal matters: perhaps he viewed the situation as one in which he must not venture guesses lest he be judged foolish.

While the above discussion is consistent with Wechsler's (1950) description of intelligence as a function of the whole personality, it is not consonant with his recommendation of factor analysis as a preferred technique for identifying components of intelligence. The inadequacy of this recommendation is that statistical techniques necessarily preclude attention to individual frames of reference.

Perhaps, if the above view proves fruitful, new intelligence tests may go beyond testing for knowledge of content and presence of "thinking processes" to include direct sampling of approach. Already there is a tendency to view creativity as an integral aspect of intelligence, and creativity may be seen as use of unusual, yet effective, approaches.

Culture

The definition's reference to competence highly regarded by the culture serves as a reminder that tests designed to measure intelligence are historically and culturally rooted. Test content reflects the dominant values of the test maker's socioeconomic class at the time of the test's construction. Thus, standard intelligence tests have stressed school-related skills. Now, in an age of computer technology, concern with remaining above mechanization has led to a renewed emphasis on creativity as a uniquely human aspect of intelligence.

As it becomes more apparent that maturity and self-esteem are necessary for the maintenance of nonviolent, nonsuppressive interpersonal relationships, these too may be directly represented on intelligence tests. Relevant here is Wesman's (1968) argument for periodic revision of both tests and school cur-

riculum in order to represent a wider spectrum of the abilities the culture has come to value. The present emphasis on culture is also consistent with the fact that scores decline with age. The decline may reflect increased specialization at the cost of breadth. Furthermore, maturity, which is associated with effectiveness of approach, is not directly represented in standard intelligence tests. Thus, many of the sources of adult effectiveness are not sampled by these tests.

Moreover, it is apparent that by definition there cannot be a culture- or class-fair test. Instead of continuing efforts toward approximating such a test, it would be better to recognize that general intelligence tests indicate a person's standing in his culture at large, regardless of the reasons. In this sense, these tests are valid for the culturally deprived in that low scores indicate low competitive standing. However, if the score is not to be a permanent label of low effectiveness, it must be accompanied by modifiers representing the tester's evaluation of the bases of the low scores, for example, "hereditary retarded intelligence," or "low intelligence in a culturally deprived child." Ideally, further data should describe the context in which the scores were earned: the testee's approaches and their degree of effectiveness, as well as other relevant aspects of the testing situation, including the examiner's role. This goes beyond the standard recommendation of stating whether an IQ is valid; it recommends a description of the ways in which the score is valid. In this manner, scores are more clearly differentiated for research as well as theorizing purposes. Further, they are available for guidance beyond mere placement.

A related asset of the definition's emphasis on the role of culture in intelligence test construction is that the appropriateness of other types of tests becomes more apparent. At times, a person's overall standing in culturally valued areas is irrelevant; for example, tests of specific abilities are more appropriate for determining school readiness, job training, or relative standing in a subculture.

Finally, thematizing the role of culture has implications for the debated existence of a general factor underlying all intelligence. To the extent that competencies valued by a

given culture at the time of the test's construction are closely related, "g" can be demonstrated through particular statistical techniques.

EFFECTIVENESS OF APPROACHES AS A CLINICAL CONCEPT

It is hoped that the proposed definition will facilitate emergent, more comprehensive conceptions and theories of intelligence. The purpose of the present section is to indicate the direction of one such conception as well as to sketch its applicability to both theory and clinical practice. The conception is holistic and existential. Specifically, it is based on the assumption that the only reality man can know is one that he co-constitutes. That is, insofar as man experiences things and events, he apprehends them through his own network of categories and meanings—his frames of reference. In turn, his behavior expresses his organizations of meaning. Frames of reference, of course, are not necessarily cognitive or logical, but are rooted in man's entire being, including its bodily aspect. This emphasis on frames of reference places the focus on the *individual* who is demonstrating intelligent behavior, rather than on abstract, disconnected processes. Such a shift of focus implies that human processes may not be as automatic (and hence logical) as present research models would suggest.

While the present position acknowledges the importance of neurophysiological structures and functions for the development of frames of reference, it asserts that the latter may not be reduced to the former. It is also recognized that frames of reference are learned and that it is important to discover how such learning occurs. Such discovery, however, will remain in terms of statistical probability until Ss are viewed as individuals. Some persons respond in a manner different from the majority because they perceive the material or event differently. The efficacy of obtaining S's experience of the stimulus situation has been empirically demonstrated by Giorgi (1967; Giorgi & Colaizzi, 1966).

The above perspective has concrete relevance for clinical use. In order to understand past behavior or predict future action, the perceptual and behavioral possibilities avail-

able to the individual must be known. In other words, a man's approaches are the limits of his possibilities, since he can act only in accordance with the meaning a situation has for him.

The method best suited for comprehending a testee's frames of reference, and thus of his present possibilities, would seem to be a descriptive one. That is, behavior-in-context should be observed, and its major trends described as accurately as possible, prior to any theoretical or explanatory activities. For example, the statement, "When required to answer verbally, the testee never altered or corrected a response," remains truer to whatever frames of reference it reflects than would reductive allusions to anxiety, overconfidence, motivation, or lack of autocriticism. Further, through such a description, a reader can view the testee more directly than he could through an examiner's derived constructs. It bears acknowledging, however, that the examiner's and reader's own frames of reference are the limiting conditions of their understanding of the testee.

Descriptive analysis is dependent on the examiner's understanding of which approaches are required for success on each subtest and each item. A form of "testing the limits," as used in projective techniques, seems appropriate for identifying absent frames of reference. Glasser and Zimmerman (1967) have shed some light on the attitudes and behaviors necessary for success on various WISC subtests. Earlier publications on the clinical usefulness of individual intelligence tests have specified item requirements only as vehicles for the demonstration of psychoanalytic constructs (e.g., Fromm, Hartmann, & Marschak, 1957; Rapaport, Gill, & Schafer, 1946).

Among intelligence tests, the Wechsler scales seem most suited to descriptive analyses, especially since the subtests require various extended relationships with the examiner. Further, they also tap different modes of relating available to the testee; for example, some subtests provide relatively clear standards which the testee can see for himself (e.g., Block Design), while others require that the testee act according to personal standards (e.g., Comprehension). Finally, many items often tap highly personal meanings of com-

mon objects and events. Together, these features afford a sample of the testee's approaches and their effectiveness. While projective techniques are better suited for ascertaining the richness of S's frames of reference, an individual intelligence test is appropriate for gauging the effectiveness of concordant behavior. The more restricted the testee's available frames of reference, the more severe is his disadvantage in our culture. Obviously, a testee's available approaches as well as effectiveness are best described through data from many sources.

The following excerpts from an intelligence test report concretely exemplify the descriptive method. The account first provided identifying information, reason for referral, a list of supplementary tests administered, and WISC scores accompanied by a statement of IQ and its statistical meaning. This was followed by a paragraph captioned "Context of Earned Scores," in which the testee's general appearance, attitudes toward the examiner, and concerns with being tested were described. Then, along with other comments related to the testee's uneven performance, the following statements were made:

When she was uncertain of what was expected of her, Diane's diligent manner gave way to sarcasm, melodrama, and rhetorical comment such as, "You don't expect me to get *THAT* do you?!!" On the other hand, when Diane could see the full parameters of a task, and knew that she possessed the required tools, she worked steadfastly, quietly, and effectively. For example, on Mazes she exceeded the scoring scale.

Quoted after this were supportive remarks made by the testee about her school work. Finally, in the Conclusions section of the report, Diane's present standing in her classroom was estimated, and it was suggested that:

Diane's growth and development will be facilitated most if her teachers understand that her occasional professed disdain for a particular activity is probably an indication that brief individual instruction is in order. That is, at such points she would profit from an explicit demonstration of what is required and how it is done, an opportunity to try her skill, and finally, reassurance that she is on the right track.

It was intended that readers of the above report would not only know Diane's levels of effectiveness, but also something about the approaches (frames of reference and their be-

havioral aspects) through which she reaches these various levels. The primarily descriptive method obviated the usual recourse to causal constructs and personality traits along with their multifarious meanings to readers. Instead, it could become apparent that for Diane, the meanings of unfamiliar tasks and uncertain expectancies are ones of danger, of something to be avoided. From intratest and intertest data, as well as interview material, these meanings emerged as a major network in terms of which she views much of her environment. Most often, this frame of reference precludes efficient problem solving. Perhaps, if her teachers follow the report's suggestion, Diane can come to see the unfamiliar in a new light, one less fraught with negative connotations.

Returning to a theoretical discussion, it is acknowledged that since frames of reference are not directly observable, their ontological status is that of a construct. In the absence of strict S-R identifications, however, they are not intervening variables. Neither are they hypothesized to exist as entities, as causal forces or energies, nor as cybernetic processes. Instead, frames of reference are proposed as abstractions of man's essential character when he is emphasized as a being in relation to others, objects, and events. That is, man is always oriented at some level to something, his frames of reference being a term for this possibility. Specific frames of reference are inferred through the question: What meanings must be experienced by the person that such-and-such a behavior is possible for him? Finally, this construct is not intended as a reduction of other phenomena, nor is it to be imposed on behavior prior to a description of the behavior itself. Justification for the last statement is use of the descriptive method, which is based on the following view of causality.

Any activity, or approach, valued enough by society to be called for on an intelligence test always occurs in the company of other necessary components (relation with tester, etc.). Since all are necessary, no one component has the status of an explanatory construct (anxiety, drive, etc.); without any one of the components, the situation does not exist. On the other hand, if the total situation

could be described, then the "if-then" requirements of explanation would be fulfilled: *if* all these components are present, *then* such-and-such *is* occurring. These components fall into two stages: in the first are those that can be abstracted as the frame of reference; in the second are those that can be seen in the activity itself.

In his article presenting an abilities conception of personality, Wallace (1966) suggested that if all of an individual's abilities and ability deficits were known, as well as the antecedent conditions necessary for their demonstration, then motivational inferences would be superfluous. To go further, it is posited here that complete knowledge of an individual's approaches can serve to indicate his available frames of reference, and hence which activities are possible for him. Explanatory constructs would, then, be unnecessary since an explanation is already constituted by a complete description of necessary and sufficient conditions. Specifically, we have bypassed the need to posit interference of one construct (intelligence) by another construct (anxiety, impulsivity, etc.).

For example, in the following statement anxiety assumes its usual explanatory status: "The boy was so anxious that he impulsively gave only partially correct responses." A descriptive analysis, however, might reveal that something like the following were necessary aspects of the boy's frames of reference: he may experience the test as something he must do well, while being uncertain of his ability, and see no effective way to stick with the test items (or, see his "impulsive" response as the best he can do).

Elements essential to activity in addition to a particular frame of reference are better known and are merely suggested here: presence of the item and physical tools, prior practice, unlimited time, feedback from the examiner, adequate vision, etc.

The level of possibility opened by a particular frame of reference should also be specified. For example, physical retaliation may have a place in a testee's network of meanings, but in what way? Perhaps its focal meaning is one of being attacked. Or, it may be that while he is aware of instances of physical retaliation around him and, at times,

can fantasy himself as an avenger, it is literally inconceivable to the testee that he might actually wreak physical vengeance on a real person. Or, the interpersonal meaning of being with another (the tester) may preclude mention of conceivable violence. Hence, if a test item asks what he should do to protect himself from an armed assailant, none of the above frames of reference would open the way to a response involving aggressive self-protection.

From the descriptive approach, more than just the testee's level of intelligence may be reported. At least some of the ways he viewed the test items can be stated. These networks of meanings, when adequately submitted to "testing of limits," may be understood as the current limits of his possible approaches. A starting point is thus provided for both retrospective investigations of how a testee's approaches came to be and predictive efforts toward specifying the conditions under which the same approaches would again be evident. These data are, of course, relevant for teaching the therapeutic efforts.

REFERENCES

- BORING, E. G. Intelligence as the tests test it. *The New Republic*, 1923, 34, 35-36.
- BOUCHARD, T. J. Current conceptions of intelligence and their implications for assessment. In P. McReynolds (Ed.), *Advances in psychological assessment*. Vol. 1. Palo Alto, Calif.: Science and Behavior Books, 1968.
- FROMM, E., HARTMANN, L. D., & MARSCHAK, M. Children's intelligence tests as a measure of dynamic personality functioning. *American Journal of Orthopsychiatry*, 1957, 27, 134-144.
- GIORGI, A. A phenomenological approach to the problems of meaning and serial learning. *Review of Existential Psychology and Psychiatry*, 1967, 7, 106-118.
- GIORGI, A. P., & COLAIZZI, P. F. Completion of simple geometric forms as a function of varied degrees of incomplete presentation. *Perceptual and Motor Skills*, 1966, 23, 1304-1309.
- GLASSER, A. J., & ZIMMERMAN, I. L. *Clinical interpretation of the Wechsler Intelligence Scale for Children*. New York: Grune & Stratton, 1967.
- GUILFORD, J. P. *The nature of human intelligence*. New York: McGraw-Hill, 1967.
- HAYES, K. J. Genes, drives, and intellect. *Psychological Reports*, 1962, 10, 299-342.
- HUNT, J. McV. *Intelligence and experience*. New York: Ronald Press, 1961.
- LIVERANT, S. Intelligence: A concept in need of re-examination. *Journal of Consulting Psychology*, 1960, 24, 101-110.
- LYONS, J. *Psychology and the measure of man*. New York: Free Press, 1963.
- MISCHEL, W. *Personality and assessment*. New York: Wiley, 1968.
- PIAGET, J. *The psychology of intelligence*. New York: Harcourt & Brace, 1950.
- RAPAPORT, D., GILL, M., & SCHAFER, R. *Diagnostic psychological testing*. Vol. 1. Chicago: Year Book Publishers, 1946.
- REITMAN, W. R. *Cognition and thought*. New York: Wiley, 1965.
- SPIKER, C. C., & McCANDLESS, B. R. The concept of intelligence and the philosophy of science. *Psychological Review*, 1954, 61, 255-266.
- TUDDENHAM, R. The nature and measurement of intelligence. In L. Postman (Ed.), *Psychology in the making: Histories of selected research problems*. New York: Knopf, 1963.
- WALLACE, J. An abilities conception of personality: Some implications for personality measurement. *American Psychologist*, 1966, 21, 132-138.
- WECHSLER, D. Cognitive, conative, and non-intellectual intelligence. *American Psychologist*, 1950, 5, 78-83.
- WESMAN, A. G. Intelligent testing. *American Psychologist*, 1968, 23, 267-274.

(Received September 23, 1968)

OPERANT-INTERPERSONAL TREATMENT FOR MARITAL DISCORD¹

RICHARD B. STUART

University of Michigan

Operant-interpersonal treatment of marital discord is premised on the assumption that successful marriages can be differentiated from unsuccessful marriages by the frequency and range of reciprocal positive reinforcements exchanged by both partners. Beginning with the clarification of behavior change objectives for each partner, a four-step treatment approach culminating in an exchange of positive responses on a reciprocal basis is suggested. The treatment of four couples complaining of low-rate conversational and sexual behavior is summarized, stressing the use of a token system as a prosthesis to facilitate the transition to increased positive reinforcement.

The operant-interpersonal approach to marital treatment rests on three assumptions concerning the character of marital interaction. First, it is assumed that the exact pattern of interaction which takes place between spouses at any point in time is the most rewarding of all of the available alternatives. This implies that the interaction between spouses is never accidental; it represents the best balance which each can achieve between individual and mutual rewards and costs (Thibaut & Kelley, 1959, p. 12). Thus when a husband consistently fails to leave his friends in order to spend time with his wife, it may be concluded that his friends offer greater relative rewards than his wife.

The second assumption is that while the specifics may vary for each couple, most married adults expect to enjoy reciprocal relations with their partners. Reciprocity has the general sociological connotation that "each party has rights and duties [Gouldner, 1960, p. 169]" and the specific behavioral connotation that each party to an interaction should dispense social reinforcement at an equitable rate (Patterson & Reid, 1967, p. 1). In effect, a quid pro quo or "something for something" arrangement underlies successful marriage (Jackson, 1965, p. 591). The exchange of rewards in marriage may be viewed as a

quasi-legal contract affording distinct safeguards to each partner. Whenever one partner to a reciprocal interaction unilaterally rewards the other, he does so with the confidence that he will be compensated in kind in the future. For example, if the husband agrees to entertain his wife's parents for a weekend, he does so with the expectation that his wife will accompany him on a weekend fishing trip at some time in the future.

Reciprocity develops as a consequence of a history of positive reinforcement. There is extensive empirical support for the proposition that ego will be more attracted to alter and will reinforce alter more if he has been positively reinforced by alter (Bachrach, Candland, & Gibson, 1961; Brewer & Brewer, 1968; Byrne, 1961, 1962; Byrne & Nelson, 1965; Homans, 1961; Komorita, Sheposh, & Braver, 1968; Newcomb, 1955; Pruitt, 1968).² When disordered marriages are evaluated in light of this reinforcement-attraction hypothesis, it is seen that each partner reinforces the other at a low rate and each is therefore relatively unattractive to and unreinforced by the other.

² Byrne and Rhamey (1965) have postulated a law of attraction magnitude which takes the following form:

$$A_x = m \left(\frac{\Sigma (PR_x \times M)}{\Sigma (PR_x \times M) + \Sigma (NR_x \times M)} \right) + k,$$

... [where] X is a positive linear function of the sum of the weighted positive reinforcements (Number \times Magnitude) received from X divided by the total number of weighted positive and negative reinforcements received from X [p. 887].

¹ Paper presented at the meeting of the Association for Advancement of the Behavioral Therapies, San Francisco, August 1968.

² Requests for reprints should be sent to Richard B. Stuart, University of Michigan, School of Social Work, 1065 Frieze Building, Ann Arbor, Michigan 48104.

The third assumption is that in order to modify an unsuccessful marital interaction, it is essential to develop the power of each partner to mediate rewards for the other. In support of this assumption, it can be shown that individuals will be more positively attracted to each other when each has been successful in influencing the other to comply with his wishes (Thibaut & Kelley, 1959, p. 124; Thibaut & Riecken, 1955). Conversely, it can be shown that when one party to an interaction fails in his influence attempts, he becomes "socially bankrupt" (Longabaugh, Eldred, Bell, & Sherman, 1966, p. 87) as he lacks the resources needed to control the other's behavior; failing to gain control through positive strategies, he resorts to negative means of control.

In successful marriage, both partners work to maximize mutual reward while minimizing individual costs. A reciprocal exchange of potent social reinforcement is established in which each partner controls sufficient rewards to compensate the other for the rewards which are expected or received from him. In an unsuccessful marriage, both partners appear to work to minimize individual costs with little apparent expectation of mutual reward. In an effort to trim costs, few positive rewards are dispensed; positive reinforcement, as a strategy of behavioral control, is replaced by negative reinforcement (removal of an aversive event following the expected response).

Either or both of two broad patterns of behavioral control, coercion and withdrawal, are likely to emerge in unsuccessful marriages. In coercion, one member seeks to gain positive reinforcement from the other in exchange for negative reinforcement (Patterson & Reid, 1967). As an example, a husband might wish his wife to express greater affection; following the failure of his amorous advances, he might become abusive, accusing his wife of anything from indifference to frigidity, abating his criticism when he receives the desired affection. The three flaws in this approach are: first, to the extent that he makes himself unpleasant, he is less likely to receive affection; second, to the extent that he is abusive or accusing, he debases his wife's affection and simultaneously reduces its reinforcing proper-

ties for himself; and third, to the extent that her affection is offered in compliance to his demand, it will appear to be appeasement rather than a gesture of genuine affection (Haley, 1963, 1967).

The withdrawal which is likely to occur in unsuccessful marriages is analogous to one of several strategies available in a prisoner's dilemma game:

If the structure of the situation is such that (a) the reward is small, (b) terminating the interaction is made difficult, and (c) the retaliatory response must take a form which is identical to the disliked other's harmful behavior, then cognitive consistency principles would lead one to expect an increase in ill will, but no retaliatory behavior [Swingle, 1966, p. 270].

Retaliatory behavior might require the husband to match his wife's actions; but to the extent that he devalues his wife for so behaving, consistency would demand that he behave otherwise. Withdrawal has the advantage of denying satisfaction to his wife while at the same time creating a situation requiring her to continue to behave assertively. Thus it is a low-risk tactic of control. The reinforcement for approach behavior on the wife's part would be termination of the husband's withdrawal. At the same time that the husband is withdrawn from his wife, however, he may also find other social and nonsocial reinforcers in his cronies, mistress, or can of beer.

Based on this formulation, the operant-interpersonal approach seeks to construct a situation in which the frequency and intensity of mutual positive reinforcement is increased. The effect of positive reinforcement in inducing positively "biased scanning [Janis & Gilmore, 1965]," or searching for assets rather than liabilities in the spouse, has been well demonstrated, particularly when the positive reinforcement is large and is offered under positive sponsorship (Elms & Janis, 1965, p. 53). It is anticipated that this positive scanning will replace negatively biased attitudes, making positive responding more likely. Positive responses, in turn, are intended to augment the range and importance of social reinforcement mediated by each spouse for the other, leading to reestablishment of successful interaction patterns.

TREATMENT CONSIDERATIONS

Operant-interpersonal treatment occurs in four orderly steps. The first step requires training the couple in the logic of the approach, consisting of two self-evident premises. The first premise is that the impressions which each spouse forms of the other is based on the behavior of the other. Accordingly, when one changes his behavior, there are corresponding changes in the other's impressions and expectations of him. Spouses who lose sight of this typically attribute marital difficulty to the personality of the other. For example, the wife who believes that her husband "is passive" implicitly suggests that her husband's personality must change before the marital disturbance can be overcome. Conversely, the wife who believes that her husband "behaves passively" need only find ways to modify her behavior (as she is in control of his actions) in order to modify his problematic responses.

This leads to the second premise, which asserts that in order to change interaction in a marriage, each partner must assume initiative in changing his own behavior before changes can be expected in his spouse. The typical couple is "locked into" problematic patterns of interaction as long as each requires a change in the other prior to changing his own behavior. If coercion and withdrawal are in fact basic problematic processes in discordant marriages, then spontaneous behavior change is highly unlikely to occur, as the response cost is too great and the potential reward too small.

Clients benefit from an explanation of the logic of the approach for at least two reasons. First, such explanations may help to free each spouse from his inaccurate and negatively biasing prejudices. Second, when each spouse is fully aware of the logic of the treatment, he can participate more fully in effective therapeutic planning and execution (Stuart, 1967).

The second step in treatment consists of asking each spouse to list the three behaviors which he would most like to accelerate in the other. This task is often subject to four difficulties. First, couples tend to begin by listing requests for decelerating negative behav-

iors, which may be expected in view of the fact that much of their interaction prior to seeking treatment has been negative. The difficulty with attempting to decelerate behavior is that its attainment would require the use of aversive stimuli or extinction paradigms, and the typical unhappy couple is already disproportionately committed to these negative strategies. Second, each spouse is likely to phrase his "three wishes" in molar rather than molecular units, and many of these are likely to be formulated in preverbal terms. For example, a husband is likely to list the request that his wife "act more feminine." "Feminine" is a modifier which is subject to varied interpretation and which may modify a wide range of behaviors. These molar requests must be reduced to specifics which include description of the desired behavior, its rate, and the context in which it is expected to occur. Third, each spouse is likely to proclaim that the other should "know what I want—if I have to tell him, then he is not as sensitive to me as he should be." Many unhappily married couples share the naïve expectation that their spouse should be clairvoyant, and it is therefore often necessary to stress the need that each must communicate his wishes to the other in order to increase the probability that they will be gratified. The fourth obstacle in this apparently simple listing of behavior change objectives is concerned with the punctuation of the behavioral chains. It has been observed that each spouse is likely to describe each unit of his own behavior as sandwiched between two negative actions on the part of his spouse (Watzlawick, Beavin, & Jackson, 1967, pp. 54-59). These three element chains (other-negative, self-positive, other-negative) obscure the true interaction, as the speaker omits reference to his stimulation of the other's initial negative actions.

The third step in the treatment calls for transcription of the three wishes of each as headings on a Behavior Checklist which is posted at some convenient point in the house. Each spouse is asked to record the frequency with which the other performs the act which he desires. While monitoring undoubtedly influences the rate of occurrence of each set

of responses, this exercise does provide a crude base line against which to evaluate change.

The fourth step consists of working out a series of exchanges of desired behaviors. The typical couple complains of a "lack of communication," which is a euphemism for a failure to reinforce each other. On closer analysis, this complaint is frequently seen as a reference to low-rate conversational and sexual behavior. This communication gap is overcome in the fourth step by arranging for each partner to compensate the other for the behaviors which he identified as socially reinforcing to him.

In marriages which have not dissolved reciprocity into coercion or withdrawal, a simple exchange of behaviors is effective. Couples are asked to accelerate desired behaviors on an equal basis. For example, one husband complained that his wife failed to greet him at the door, that she did not straighten the family room in anticipation of his return home from work, and that meals were rarely ready on time. His wife complained that he failed to spend sufficient time with the children (it was agreed that 30 minutes before bed was sufficient), failed to take her out for an occasional movie, and failed to pay attention to meals when they were well prepared. Accordingly, each of these behaviors was restated as a positive (e.g., greet husband at door) and was listed on the Behavior Monitoring Form. Each person recorded the frequency with which the other completed the desired behavior. When these behaviors were accelerated at a sufficient rate, other goals were added.

In marriages in which reciprocity appears to be essentially absent, a token economy has proven useful. Where reciprocity fails, couples tend not to trust each other. It is therefore important to arrange for some immediate form of reinforcement. Tokens are an ideal media for such reinforcement because (a) they are given immediately, (b) they can be redeemed for the specific consequences which the recipient deems desirable at that point in time, (c) they are concrete and unambiguous, (d) the giving and receiving of tokens is customarily associated with positive social

interchange, and (e) they permit an exchange of behaviors which are not contiguous. Token systems have been used effectively in a wide range of settings (e.g., mental hospitals—Atthowe & Krasner, 1968; classrooms—Clark, Lachowicz, & Wolf, 1968; and institutions for delinquents—Tyler & Brown, 1967), and extension of this technique to marital treatment appears warranted.

ILLUSTRATION AND RESULTS

Four couples have used the token system to modify each other's behavior. Individuals ranged in age from 24 to 52 and in education from high school diploma to doctorate. The couples were married for from 3 to 23 years and had a maximum of two children. Each of the couples sought treatment as a last-ditch effort prior to obtaining a divorce. In each instance, the wife listed as her first wish that her husband converse with her more fully, or at least that he not "close me out of his life even when he is at home." Considerable discussion was often necessary to identify what intensity level of conversation was positively reinforcing to the wife, and this was made clear and rehearsed during the treatment sessions. The wife was then instructed to purchase a kitchen timer which she could carry with her about the house. She was instructed to set the timer as soon as her husband entered and to give him one token when the bell rang after each hour in which he conversed at the criterion level. If he failed to behave at the criterion level by the end of the first 30 minutes of each hour, she had to notify him of this and offer constructive suggestions, cueing him as to how his performance could be improved upon. If she failed to do this, he had to be given a token even if he failed to perform adequately. If he so requested, at the half-hour cue time, the timer could be reset so that he could earn a token during the next hour (so that he waited 60 rather than 90 minutes before being rewarded).

The criterion level for conversation is naturally a negotiable factor. No one could be expected to talk to his wife constantly; if for no other reason, his children would not allow it. Therefore, conversational tokens

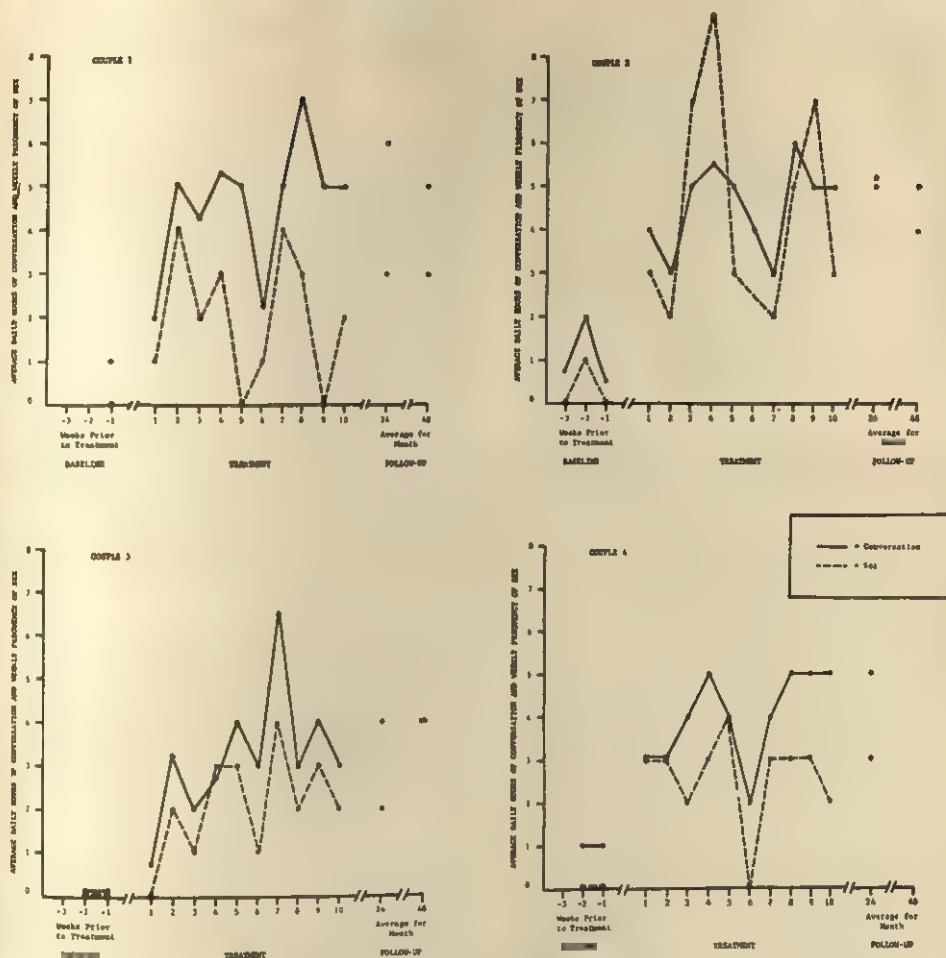


FIG. 1. Average daily hours of conversation and weekly rate of sex of four couples—before, during, and after operant marital therapy.

can be earned for a wider range of responses ranging from intense conversation at one extreme to the wife's feeling free to interrupt her husband with a question at agreed intervals at the other extreme.

While tokens may have some intrinsic reinforcing properties in their own right (in addition to being associated with positive social responses when they are offered), they become more powerful when they function as contingencies for some other event. With the four couples cited, tokens were redeemable at the husband's request from a menu stressing physical affection. A different menu was constructed for each couple which took into account the base-line level of sexual activity, the desired level of sexual activity, and

the number of hours available for nonsexual (in this instance conversational) interchange. Each of these couples had sex less than once per week (ranging from once in the year prior to treatment to once in the week prior to treatment), each desired sex an average of three times per week, and each had approximately 5 hours together on weeknights and 14 hours on weekends, making a total of approximately 52–54 hours per week. Accordingly, husbands were charged three tokens for kissing and "lightly petting" with their wives, five tokens for "heavy petting," and 15 tokens for intercourse. (These behaviors were not rehearsed during treatment sessions.)

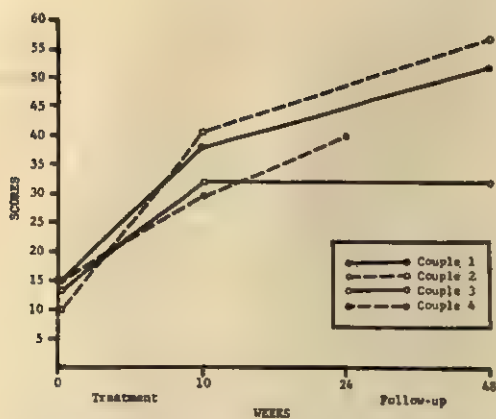


FIG. 2. Marital satisfaction assessment inventory scores of four couples—before, at last session, and at follow-up of operant marital treatment.

Tokens earned and spent were recorded on the Behavior Checklist, which also provided data for continued graphing of interactional behavior. The performance of each of these couples is represented in Figure 1, where it will be seen that the rates of conversation and sex increased sharply after the start of treatment and continued through 24- and 48-week follow-up periods.

At the start of treatment, at the conclusion of regularly scheduled interviews, and at the time of follow-up, each spouse was asked to complete a brief inventory measuring the extent of his own and his perception of his spouse's satisfaction in and commitment to the marriage. This inventory was adapted from the work of Farber (1957). The results are depicted in Figure 2 where it can be seen that the rate of reported satisfaction increased in association with the reported behavioral changes. These changes enabled the couples to become more similar to nonclinic families on this dimension (Levinger & Breedlove, 1966, p. 369).

With each of these couples, all therapeutic sessions were held jointly. Sessions were held during the first four, the sixth, eighth, and tenth weeks, for a total of seven sessions. When it is considered that these couples were each on the brink of filing for divorce, this could be considered relatively inexpensive treatment. Follow-up contacts were held by phone or by mail, and all data, including

that collected during sessions, were based on self-report.

DISCUSSION

The therapeutic recommendations of alternative courses of behavior are consistent with Goldiamond's (1966, p. 118) recommendation to his patients that: "If he wished his wife to behave differently to him, then he should provide other stimuli than the ones which produced the behaviors he did not like." Each spouse was directed in specific modifications of his own behavior in an effort to modify the behavioral environment in which his partner's behavior occurred. The antecedents for changed behavior were probably twofold: therapeutic recommendations exercised some discriminant stimulus control as did the expectation of changes in the behavior of the spouse. The therapist suggested that each party engage in behaviors which had doubtless been requested, cajoled, and demanded by each party countless times before. As therapeutic directives, these new requests were differentiated from the old in two important respects. First, they were clarified. Second, they were removed from the context of coercive demands in which granting the request would have been tantamount to reinforcing dysfunctional patterns of negative reinforcement. In addition, the treatment was characterized to each couple as a "game" in which each would be able to modify the general rules of the relationship so that positive rather than negative strategies would pay off.

It is impossible to isolate which aspects of this complex treatment approach carry the weight of the observed change. Behavioral rehearsal and the manipulation of social reinforcers were the objectively identified variables but other observers would undoubtedly stress other factors as well. One technical point of particular importance is the fact that at no time was an attempt made to "fade" (or slowly remove) the behavior modification system, replacing it with more natural processes of behavioral control (Krasner, 1965, 1966). Indeed, continued reliance on the therapeutic techniques as a means of programming and evaluating each other's behavior could result in considerable gain with low

cost. In fact, one characteristic of an ideal therapeutic technique might be that it is transferable to the natural environment without change.

This pattern of treatment is likely to be challenged on two counts. Some critics will charge that the therapeutic strategies of this approach are "trivial" because they are based on "superficial" changes in behavior. Two counterarguments should be stressed. First, the only data which are available to the therapist as a scientist are observable behaviors, and it is at the level of behavior that changes must be sought (Skinner, 1953). Second, it must be stressed that these were the changes sought by the clients and were therefore in accord with the therapeutic contract. To have sought other goals would have meant ignoring the presenting complaints.

In response to the triviality argument, it is also relevant to cite the data concerning the level of marital satisfaction. However, these must be considered to be "soft data" as they are based on global self-ratings such as "How committed are you to stay in this marriage?" and "What proportion of time spent with your spouse would you consider to be 'fully satisfying'?" It is impossible to determine exactly what degree of meaningfulness can be attributed to answers to such questions and they should, at best, be interpreted as indications of trend.

REFERENCES

- ATTHOWE, J. M., JR., & KRASNER, L. Preliminary report on the application of contingent reinforcement procedures (token economy) on a "chronic" psychiatric ward. *Journal of Abnormal Psychology*, 1968, 73, 37-43.
- BACHRACH, A. J., CANDLAND, D. K., & GIBSON, J. T. Group reinforcement of individual response experiments in verbal behavior. In I. A. Berg & B. M. Bass (Eds.), *Conformity and deviation*. New York: Harper, 1961.
- BREWER, R. E., & BREWER, M. B. Attraction and accuracy of perception in dyads. *Journal of Personality and Social Psychology*, 1968, 8, 188-193.
- BYRNE, D. Interpersonal attraction and attitude similarity. *Journal of Abnormal and Social Psychology*, 1961, 62, 713-715.
- BYRNE, D. Response to attitude similarity-dissimilarity as a function of affiliation need. *Journal of Personality*, 1962, 30, 164-177.
- BYRNE, D., & NELSON, D. Attraction as a linear function of proportion of positive reinforcements. *Journal of Personality and Social Psychology*, 1965, 1, 659-663.
- BYRNE, D., & RHAMEY, R. Magnitude of positive and negative reinforcements as a determinant of attraction. *Journal of Personality and Social Psychology*, 1965, 2, 884-889.
- CLARK, M., LACHOWICZ, J., & WOLF, M. A pilot basic education program for school dropouts incorporating a token reinforcement system. *Behaviour Research and Therapy*, 1968, 6, 183.
- ELMS, A. C., & JANTS, I. L. Counter-norm attitudes induced by consonant versus dissonant conditions of role-playing. *Journal of Experimental Research in Personality*, 1965, 1, 50-60.
- FARBER, B. An index of marital integration. *Sociometry*, 1957, 20, 117-134.
- GOLDIAMOND, I. Self-control procedures in personal behavior problems. In R. Ulrich, T. Stachnik, & J. Mabry (Eds.), *Control of human behavior*. Chicago: Scott, Foresman, 1966.
- GOULDNER, A. W. The norm of reciprocity: A preliminary statement. *American Sociological Review*, 1960, 25, 161-178.
- HALEY, J. *Strategies of psychotherapy*. New York: Grune & Stratton, 1963.
- HALEY, J. Marriage therapy. In H. Greenwald (Ed.), *Active psychotherapy*. New York: Atherton Press, 1967.
- HOMANS, G. C. *Social behavior: Its elementary forms*. New York: Harcourt, Brace & World, 1961.
- JACKSON, D. D. Family rules. *Archives of General Psychiatry*, 1965, 12, 589-594.
- JANIS, I., & GILMORE, J. The influence of incentive conditions on the success of role playing in modifying attitudes. *Journal of Personality and Social Psychology*, 1965, 1, 17-27.
- KOMORITA, S. S., SHEPESH, J. P., & BRAVER, S. L. Power, the use of power, and cooperative choice in a two-person game. *Journal of Personality and Social Psychology*, 1968, 8, 134-142.
- KRASNER, L. Operant conditioning techniques with adults from the laboratory to "real life" behavior modification. Paper presented at the meeting of the American Psychological Association, Chicago, September 1965.
- KRASNER, L. The translation of operant conditioning procedures from the experimental laboratory to the psychotherapeutic interaction. Paper presented at the meeting of the American Psychological Association, New York, September 1966.
- LEVINGER, G., & BREEDLOVE, J. Interpersonal attraction and agreement: A study of marriage partners. *Journal of Personality and Social Psychology*, 1966, 3, 367-372.
- LONGABAUGH, R., ELDRED, S. H., BELL, N. W., & SHERMAN, L. J. The interactional world of the chronic schizophrenic patient. *Psychiatry*, 1966, 29, 78-99.
- NEWCOMBE, T. M. *Social psychology*. New York: Holt, Rinehart & Winston, 1955.

- PATTERSON, G. R., & REID, J. Reciprocity and coercion: Two facets of social systems. Paper presented at the meeting of the Institute for Research in Clinical Psychology, Lawrence, Kansas, April 1967.
- PRUITT, D. G. Reciprocity and credit building in a laboratory dyad. *Journal of Personality and Social Psychology*, 1968, 8, 143-147.
- SKINNER, B. F. *Science and human behavior*. New York: Macmillan, 1953.
- STUART, R. B. *Guide to behavior modification*. Unpublished manuscript, University of Michigan, 1967.
- SWINGLE, P. G. Effects of the emotional relationship between protagonists in a two-person game. *Journal of Personality and Social Psychology*, 1966, 4, 270-279.
- THIBAUT, J. W., & KELLEY, H. H. *The social psychology of groups*. New York: Wiley, 1959.
- THIBAUT, J. W., & RIECKEN, H. W. Some determinants and consequences of the perception of social causality. *Journal of Personality*, 1955, 24, 113-133.
- TYLER, V. O., JR., & BROWN, G. D. Token reinforcement of academic performance with institutionalized delinquent boys. Paper presented at the meeting of the Western Psychological Association, San Francisco, May 1967.
- WATZLAWICK, P., BEAVIN, J. H., & JACKSON, D. D. *Pragmatics of human communication: A study of interactional patterns, pathologies and paradoxes*. New York: Norton, 1967.

(Received October 3, 1968)

PERSONAL ADJUSTMENT AS A DETERMINANT OF AGGRESSIVE BEHAVIOR TOWARD THE MENTALLY ILL¹

KENNETH RING AND AMERIGO FARINA²

University of Connecticut

Thirty-six Ss participated individually in an experiment which required them to act as teachers whose job was to enable a person whom they believed to be an ex-mental patient to learn an arbitrarily correct series of button presses. The Ss could only "communicate" with the learner by shocking him when he was incorrect and could vary both the intensity and duration of the shock. It was found, using mean shock duration as an index of punitiveness, that the more emotionally unstable (and thus similar to the learner) an S believed himself to be, the less punitive he was toward the learner. The surprising utility of the shock duration measure as an index of aggression against a stigmatized person is discussed.

It has been clear for some years that a person afflicted by conditions such as mental illness or blindness is stigmatized and degraded in our society (Nunnally, 1961; Nunnally & Kittross, 1958). More recently, owing partly to the stimulus provided by Goffman's (1963) book, *Stigma*, a series of experimental studies was carried out which sought to specify some of the variables affecting the perception and treatment of stigmatized individuals (Farina, Allen, & Saul, 1968; Farina, Holland, & Ring, 1966; Farina & Ring, 1965; Farina, Sherman, & Allen, 1968). The results of these studies support Goffman's belief that stigmata are important determiners of interpersonal interactions. However, most of these studies have varied (perceived) characteristics of the stigmatized individual such as the severity of the stigma and the degree to which he can be held accountable for it. In the present study, on the other hand, the focus was on the other party in the stigmatizing process, the person in interaction with the stigmatized individual. Unquestionably there are a multitude of characteristics which are related to the way a person feels and behaves toward a blemished other. The characteristic which was examined in the present study was

psychological similarity. It was systematically varied while stigma remained constant.

There is a dearth of experimentally based findings which indicate that perceived similarity to a stigmatized individual influences the way one responds to him. Nevertheless, it seems almost self-evident that such a relationship exists and there is anecdotal evidence. Thus Wright (1960; cited by Goffman, 1963) indicates that under some circumstances the hard-of-hearing person may respond more harshly than the completely deaf to one who is deaf. A striking illustration of the way in which one's similarity to a stigmatized group can increase aggressiveness toward that group has recently been given by Yablonsky (1966). He reports interviewing a white New York gang boy who was asked by a person he perceived as a Negro, but who had a Spanish accent, to join a gang that intended to kill all the "niggers." Clearly the dark-complected man did not regard himself as a Negro.

These examples imply a curvilinear relationship between similarity and punitiveness with intermediate similarity leading to greater punitiveness. It does not seem implausible that a person who is quite uncertain about the degree to which he resembles a stigmatized person, in contrast to those who know with certainty that they do or do not possess the stigma, would be the most anxious to establish his distance from the stigmatized person. This hypothesis implies that persons of un-

¹ This study was financed by Research Grant GS-848 from the National Science Foundation.

² Requests for reprints should be sent to Amerigo Farina, Department of Psychology, University of Connecticut, Storrs, Connecticut 06268.

certain or intermediate similarity to the stigmatized individual would behave in the most aggressive fashion toward the latter as a means of dissociating themselves from them. Rubin's (1967) recent finding that an increase in self-acceptance leads to a reduction in ethnic prejudice also seems consistent with this hypothesis.

An alternative and more straightforward hypothesis is that with increasing similarity there would be decreasing punishment administered by Ss. This would happen if the closer one feels to a stigmatized individual, the more one identifies with him and feels compassion for him. Milgram's (1965) experiment is relevant since it shows that decreasing psychological distance between teacher and victim in his obedience experiment results in a decrease in punitiveness.

One purpose of the present experiment was to permit a test between these two hypotheses, the curvilinear and the negatively monotonic one. The authors felt the effect of psychological similarity would be enhanced under conditions of public disclosure as is suggested by the case of the New York Grand Dragon of the Ku Klux Klan who committed suicide when he learned that his Jewish identity had been discovered. Therefore, all of the Ss were led to believe that their stigmatizing condition (emotional instability) would become a part of their official college record.

METHOD

Subjects

Thirty-six male introductory psychology students served as Ss in this experiment. Participation in psychological experiments was a requirement of the course and each S received 2 hours' credit for this experiment. The Ss were assigned randomly in equal numbers to one of three experimental conditions.

Procedure

First session. When S arrived at the experimental laboratory, he was told by E³ that the study dealt with the evaluation of rehabilitation techniques used for mental patients. The experiment, E explained, had to do with "what kind of people work best with mental patients in the sense of communicating with them and in teaching them something." The E went on to outline some (fictional) views of psychologists concerning whether persons similar to or dissimilar from mental patients are most effective in

³Thanks are due Kenneth Wallston for conducting the experiment.

this regard, and indicated that the present experiment would attempt to find out.

It was then explained that the experiment consisted of two parts which would require that he return some 4 or 5 days later. In the second part, one person would attempt to teach something to another. The teacher would always be a college student but the learner would be one of a group of former mental patients who had volunteered to participate for pay. In order to participate as a teacher, however, the student had to fall into one of three rather narrow categories of emotional stability: very stable, intermediate, or very unstable. He was told a Rorschach test would now be given and it would be scored prior to the next session. If he fell into one of the three categories, he would serve as a teacher in the subsequent session, whereas, if he did not, he would merely learn the result of the testing and be excused. The Ss were then given a chance to withdraw from the experiment and were also told that their test results would become a part of their official records unless they objected. No one withdrew and no one objected.

The first five cards of the Rorschach test were administered to all Ss. If S was assigned to the emotionally stable (ES) condition, E's manner was businesslike and nonevaluative. For those in the emotionally unstable (EU) condition, E pretended to be surprised at some of the answers and, in general, tried to give S the impression the responses were unusual. The E's behavior toward Ss in the emotionally intermediate (EI) condition fell somewhere between the two extremes just described. The purpose of E's differential behavior was, of course, to prepare S to receive the particular psychological evaluation that chance rather than his own performance had dictated for him.

Second session. When he returned for his second appointment, each S was told he had a score which qualified him for the second phase of the study, but E was kept ignorant of the specific category to minimize bias. Hence, E could answer no questions, but he could arrange an appointment with one of the two psychologists who had scored the test. The S was here given a sealed envelope and told to spend a little while looking over his evaluation. Those in the ES condition found their emotional adjustment had been rated "very satisfactory" and ranked in the top 5% of college students in emotional stability. Those in the EI condition read that there were some signs of good and some signs of poor emotional stability which we hoped would induce uncertainty concerning how emotionally unstable they really were. Those in the EU condition received the following evaluation which we present in its entirety to convey the flavor of these "reports."

Mr. _____'s test responses show quite unequivocally that his present emotional adjustment is decidedly unsatisfactory and a cause for concern. There is clear evidence from his protocols of serious unconscious factors which, when coupled with stressful external circumstances, might well

precipitate a breakdown. It is precisely because these difficulties do have an unconscious basis that they are doubly dangerous. While Mr. _____ is no doubt vaguely aware of some psychological difficulties, it is extremely unlikely that he recognizes just how tenuous his emotional adjustment really is. His responses, which were unimaginative and constricted, indicated serious difficulties both in the expression of aggression and the management of sexual tensions. In short, I should say that Mr. _____ is quite an emotionally unstable student who is likely to experience difficulties in the future. On the basis of my experience in evaluating the emotional stability of college students based on the Rorschach, I should place Mr. _____ in the bottom 5%. In addition, I should recommend that he avail himself of the psychiatric services provided by the on-campus clinic if he is not already doing so.

The three evaluations were approximately equal in length and were signed by one of the authors, with his PhD and clinical psychologist status indicated.

At this point, the student was told that he and the learner would be in different rooms to protect the ex-patient's anonymity. The task entailed teaching a pattern of five button presses. He was shown the teacher's apparatus which had three lights labeled A, B, and C, as well as a button labeled "Shock" which could be set at any of 10 positions. He was told the lights corresponded to three buttons on the learner's panel and his task was to teach the learner the pattern AABCA by shocking him each time he was wrong. He could set the shock from light to painful and administer it for as long as he wished but only 30 trials would be allowed. In the next room there was a confederate who communicated the same 30 sequences of five button presses to all Ss, approaching and later departing from the correct sequence so that on the last trial he was as far as ever from the correct pattern. The confederate also recorded for each trial the level of shock used and the duration of the shock. This procedure is very similar to that employed in earlier reports and the apparatus is more fully described there (Farina et al., 1966; Farina, Sherman, & Allen, 1968).

After the thirtieth trial, the student was stopped and it was pointed out that the learner had not performed his task very well. He was then asked to complete a postexperimental questionnaire designed to check the effectiveness of the experimental manipulation as well as to assess S's perception of the learner and evaluate the latter's and his own performance.

Finally, a full debriefing session followed which had three objectives: (a) to determine whether and to what extent Ss were suspicious about any of the deceptions employed in the experiment—nine Ss were suspicious about some aspects of the procedure and were dropped from the study; (b) to explain in detail the real purpose of the experiment and to justify its several deceptions; and (c) to allow Ss to express their feelings and make any comments

TABLE 1
NUMBER OF SUBJECTS INTERPRETING PSYCHOLOGICAL
EVALUATION ACCORDING TO EXPERIMENTAL
CONDITION

Interpretation	Experimental condition		
	ES	EI	EU
Emotionally stable	10	2	0
Somewhat stable	1	11	0
Definitely unstable	0	1	11

they cared to about the experiment. The Ss were encouraged to speak to the authors if they continued to be bothered by any aspect of the experiment. The Ss were promised a full write-up of the experiment and were, once again, before being dismissed, sworn to secrecy.

Since the general procedure of this study as well as some of the "psychological" evaluations used could be expected to upset some of the Ss, at least during the experiment itself (Ring, Wallston, & Corey, in press), the present authors felt incumbent to determine whether there were any long-term negative aftereffects. Data relevant to this issue will be presented in the Discussion section.

RESULTS

Checks on the Manipulations

Before examining other data, it is necessary to determine if Ss interpret the evaluations they received in accordance with the experimental intention. To do this a question was contained in the postexperimental questionnaire which asked Ss to indicate how emotionally stable they were according to the evaluation. The numbers choosing each of the three alternatives for each condition are shown in Table 1. The table shows quite clearly that nearly all Ss accurately characterized the type of evaluation they received. Furthermore, the students responded emotionally to these evaluations in the expected and desired direction. To determine this, they were asked to indicate how upset they were after reading the evaluations by checking the choices: 1. not at all; 2. a little; 3. fairly; 4. very much. Upset scores were derived by assigning numbers ranging from 0 for the first to 3 for the last choice. The average scores for the groups were as follows: ES = .42; EI = .83; EU = 1.50. These group scores are

TABLE 2

MEAN DURATION OF SHOCK ADMINISTERED FOR
THREE BLOCKS OF 10 TRIALS EACH AND
FOR ALL 30 TRIALS

Condition	Block			All trials
	1	2	3	
ES ¹	13.86 ^a	15.15	13.69	42.70
EI ¹	10.43	11.85	10.72	33.00
EU	8.46 ^a	8.48	8.25	25.19

significantly different ($F = 5.04$, $df = 2/33$, $p < .025$).

The other condition we wanted to create was to have Ss believe a copy of the evaluation would be sent to their counselor. To check on this, a question asked them to indicate what *E* had said concerning the report. Of the 35 Ss who answered this question, all but one indicated clearly that they remembered what *E* had told them concerning the disposition of their test reports. The remaining student wrote, "I care very little about it." It may be concluded, therefore, that the conditions necessary to test the hypothesis were satisfactorily established.

Treatment of the Stigmatized Person

As the principal measure of interpersonal aggression, mean trial shock duration was used, partly because other work using the same pseudoshock apparatus (Farina et al., 1966; Farina, Sherman, & Allen, 1968) indicates the duration measure is a more sensitive index of interpersonal aggression than either intensity alone or a multiplicative index based on intensity and duration scores, although in this experiment, all three measures reveal very much the same pattern of condition differences. Other investigators (e.g., Berkowitz, 1965; Hartmann, 1965), using similar shock devices, have also found a duration measure to discriminate significantly among experimental treatments. Additional reasons for the exclusive reliance on this type of index, based on S's phenomenology, will be given in the Discussion section.

The mean shock duration data are presented in Table 2. The mean scores for each block of 10 trials are given in the table as

well as the average score for all 30 trials. As may be seen, there is a clear experimental condition effect with decreased punitiveness evident with increasing psychological similarity to the stigmatized learner. This effect, based on raw scores, is only marginally significant ($F = 2.79$, $df = 2/33$, $.05 < p < .10$). However, because of the skewness of the duration scores, the authors followed Berkowitz (1965) and subjected them to a square root transformation ($\sqrt{x} + \sqrt{x + 1}$) which resulted in an $F = 3.68$, $p < .05$. Consistent with this finding, a Kruskal-Wallis one-way analysis of variance (based on ranks) generated a significant treatment effect ($.02 < p < .03$), with the ranking of conditions precisely the same as in the parametric analysis of variance (i.e., $ES > EI > EU$). The results of all of these analyses, then, support the conclusion that in this experiment a negatively monotonic relationship exists between psychological similarity to a stigmatized person and aggression toward him—the more one feels himself to be different from a blemished individual, the greater is one's punitiveness toward him. The trend effect was not significant indicating no differences among the blocks of trials.

Perception of the Stigmatized Person

A number of items were included on the postexperimental questionnaire which were designed to measure possible differences among conditions in either self- or other-perceptions. For the most part, however, these items did not differentiate among conditions. The authors did find that the more similar an S felt to the confederate, the most confident he was that the latter would have learned the pattern had there been more trials, but the effect was only marginally significant ($.05 < p < .10$ by an F test; $p < .05$ by a χ^2). The only other significant difference ($p < .05$ by an F test) on a perceptual variable indicated that, compared to Ss in other conditions, EU Ss believed that the learner was experiencing less pain from the shocks. This is a trivial finding, however, since it was precisely these Ss who administered less intense as well as briefer shocks. And in fact, the correlation between estimated average shock intensity and actual intensity was .93

($p < .001$) while the corresponding figure for duration was .38 ($p < .05$).

The negative result which is difficult to understand concerns the perceived similarity of the student to the patient. To measure this, four items were included in the post-experimental questionnaire which asked *S* in various ways to indicate to what extent he was psychologically similar to the learner. The scores of these four items were summed to form a distantiation index and the three groups of *Ss* were essentially identical with respect to this index. In view of the success of the manipulation, differences among the groups should have emerged. Conceivably ego-defensive needs led those *Ss* who thought they were emotionally unstable to indicate being no more similar to the former patient than those who thought they were stable. However, we have no data to substantiate this or other possible hypotheses.

Relationships between Perception and Treatment of the Stigmatized Person

The results presented thus far imply that psychological similarity has a rather systematic and straightforward effect on one's behavior toward a person believed to be an ex-mental patient but no consistent relationship to one's perception of him. From this it would appear that the treatment and evaluation of such a blemished individual do not necessarily correspond, a point that has been made before (Farina et al., 1966). One bit of evidence, however, suggests that under some circumstances there might be a relationship. For *Ss* in the EI condition, there is a correlation of .51 ($p < .10$) between felt similarity as measured by the distantiation index and duration of shock administered, whereas, for the other two conditions the correlations are $-.09$ and $-.14$. Thus, in the EI condition, the more similar an *S* feels to the ex-mental patient, the more he shocked him. These data are consistent with the curvilinear hypothesis and raise the possibility that there was something psychodynamically special about the middle group.

DISCUSSION

Except for the tenuous finding just mentioned which tends to complicate matters a

little, the results of this investigation appear fairly clear-cut. They lead to the general conclusion that a person's inclination to inflict pain on a stigmatized individual is blunted to the degree that the former believes himself to be psychologically similar to the latter. This generalization, nevertheless, should not be accepted too readily for it raises a number of issues that can only be discussed but not settled here.

First of all, in what sense can it be asserted that *Ss* in the EU condition, for example, believe themselves to be similar to the ex-mental patient? Their questionnaire responses indicate that, compared to other *Ss*, they do not perceive themselves as any more similar to the learner. Assuming that these responses are not due to a conscious ego-defensive maneuver, it would appear that the independent variable has behavioral effects which occur without any necessary corresponding cognitive representation. The lack of cognitive parallels for these behavioral differences is in itself a matter deserving of attention and will be dealt with below.

Another important qualification of the conclusion stated at the outset hinges on the dependent variable which differentiated the conditions, namely, duration of shock. Why should this measure consistently relate to treatment of the stigmatized and not intensity of shock which is perhaps the more obvious variable? The authors' interpretation rests on the assumption that there are strong normatively based restraints against treating former mental patients (and probably many classes of stigmatized persons) differently, and especially more negatively than "normal" persons. At all events, the prescriptions of American middle-class society demand that we respond to disadvantaged persons as if we were not aware of, or if it's unavoidably noticeable, as if we did not care about, their stigma. We are not, in short, supposed to hold it against a man that he has been, for example, a prisoner or a mental patient. At the same time, it is undeniable that private attitudes toward such persons are often incompatible with these normative prescriptions.

The authors submit, therefore, that we are most likely to inhibit normatively inappropriate responses (e.g., condemnation, ridicule,

physical aggression, etc.) toward a disadvantaged person when we can easily label or recognize them as such. It would be expected then that aggression toward the stigmatized would ordinarily assume more covert, indirect, or subtle form, thus keeping the aggressing individual often largely unaware of both the meaning and motivation of his actions—a state of affairs that could serve an important guilt-minimization function for the person.

In support of the foregoing argument, the data suggest that shock duration is a more subtle mode of aggression than shock intensity. The *Ss* were asked to estimate both the mean shock intensity and shock duration that they administered to the learner. When these estimates were correlated with the objective values of these variables, the relationship, it will be recalled, was nearly perfect ($r = .93$) for intensity, but only mildly positive for duration ($r = .38$). The difference between these two correlations was highly significant ($z = 5.12$, $p < .0001$). Obviously, it was much easier for *Ss* to monitor shock intensity, which they had to select before each shock, than shock duration for which they had no objective feedback.

This interpretation, emphasizing the often covert character of aggressive acts toward the stigmatized, may also help to illumine the curious finding that revealed that it was the seemingly more sympathetic EI *Ss* who administered the longer shocks. If this is in fact a reliable relationship (and we repeat that it is most conjectural and demands replication), it points to the psychodynamically appealing proposition that among those persons who feel most uncertain concerning whether they too are stigmatized, those who are verbally the most positive in their attitudes toward the stigmatized are likely to be the most covertly punitive in their acts.

Finally, the view taken here can be used to account for the somewhat puzzling finding (as in the Farina, Sherman, & Allen, 1968, study as well as the present one) that significant behavioral differences toward stigmatized persons occur in the absence of concomitant perceptual or attitudinal differences. The authors propose that since many of the questionnaire items typically used in these

studies are patently evaluative in content, *Ss* can be expected ordinarily to respond to them in normatively appropriate ways. The authors also anticipate that items designed to tap evaluative dimensions but in more covert or nonobvious ways would be more successful in eliciting differences which would parallel behavioral measures.

It is worth pointing out that the present position can easily be subjected to empirical disconfirmation. To illustrate one possibility, since shock duration is regarded as a sensitive indicator of aggression only because variations in duration are difficult for *S* to be aware of, any procedure which makes duration as a variable more salient or easy to monitor (e.g., requiring *S* to record time of shock or feedback from confederate related to shock duration), for *S* ought to attenuate or obliterate altogether duration differences among treatment conditions similar to those used in previous experiments.

So far the authors have only suggested an interpretation which accounts for the differentiating reliability of the duration measure; there has been no attempt, it should be noted, to provide an explanation for the findings of the present experiment. The question here is: Why should similarity lead to decreased punitiveness?

The authors would like to argue, as was proposed in the introduction, that perceived similarity increases the likelihood of identification which, in turn, leads to a reduction in aggression. However, there is no direct evidence to support this interpretation and, as noted earlier, it is not even clear that perceived similarity did vary among the experimental groups as it was expected to. Furthermore, there are several plausible alternative explanations that cannot be ruled out. For example, it may simply be that *Ss* who receive a psychological evaluation which deals a blow to their self-esteem are less likely to be punitive toward another person, whether he be stigmatized or not. Or it may be that when one is generally upset or preoccupied with his own feelings, one is not as likely to be punitive. It seems to the present authors that without the appropriate control groups which might eliminate these alternative interpretations, a high degree of confidence cannot be

placed in the particular explanation that is favored here. In addition to further research designed to evaluate the tenability of competing formulations, data are needed which will more directly corroborate the "identification hypothesis" than those available from the present study.

Considered now are some reasons why the curvilinear hypothesis was not supported more strongly by the data. The post hoc view is that the heightened punitiveness toward a stigmatized person that this hypothesis predicts is most likely to occur when (a) *S* is aware that he has the stigma or thinks or fears that he does and (b) he fears that important others will learn of his real or suspected condition. It is crucial to note that this revised formulation requires the presence of both conditions in order for the effect to be observed. The present experiment presumably established the first condition, but one is left in doubt concerning the second. It is not at all clear that *Ss* thought it was very likely that the evaluations would ever really fall into the hands of persons whom they would wish not to know about their condition or just how upset they would be if such information were revealed to others. They were not, after all, now or formerly mental patients—it had only been implied for some that they might be. Whether being told that one is emotionally unstable and a possible candidate for a mental institution can be regarded as a stigma (let alone a severe one) is conjectural. What is needed to test this hypothesis definitively would seem to be an experiment which varies both an *S*'s motivation to conceal the stigma and the probability that it will remain concealed. Until such an experiment is conducted, it would be premature to abandon the (now revised) curvilinear hypothesis.

One more point unrelated to any of the substantive or conceptual issues raised by this experiment requires brief mention. It was asserted earlier that certain ethical questions could be raised concerning the procedures used in this study. While the following results do not in themselves resolve all such doubts about this kind of research, they do answer some of the more obvious questions

dealing with the long-term impact of this particular experiment.

Some weeks after the conclusion of the study, open-ended questionnaires, to be filled out anonymously, were sent to each *S*. These questionnaires attempted to elicit an *S*'s retrospective account of his feelings during the experiment, immediately after it was over, as well as his present thoughts and feelings about it. Sixteen questionnaires were returned. One might think that if any selective factor were operative, it would favor those *Ss* who were most upset by *E* returning their questionnaires.

The results, in any case, were overwhelmingly negative. Only one *S* expressed any regret about having been in the experiment and that was because he wasn't certain he had been told the truth. Most *Ss* actually said that they had enjoyed participating in the experiment and that they hadn't at all minded being deceived. These findings strongly support Milgram's (1964) views about the lack of serious negative long-term effects from severe deception experiments and are also corroborated by the results of a more systematic experiment (Ring et al., in press) specifically directed to the question of immediate and long-term consequences of being an *S* in a Milgram-type obedience experiment. The findings from that experiment, too, offer no support to the position expressed by some (e.g., Baumrind, 1964) and no doubt feared by many others that there are likely to be worrisome long-term consequences for many *Ss* participating in such experiments.

The only negative effect inferable from these questionnaire data stems from the statements of a couple of *Ss* that they thought they might be more suspicious about psychological experiments in the future. In view of some recent findings about the effects of deception experiments on subsequent experimental performance (e.g., Brock & Becker, 1966; Fillenbaum, 1966), it is not immediately obvious that these suspicions about future suspiciousness should in themselves be regarded as particularly alarming.

REFERENCES

- BAUMRIND, D. Some thoughts on ethics of research: After reading Milgram's "Behavioral Study of

- Obedience." *American Psychologist*, 1964, 19, 421-423.
- BERKOWITZ, L. Some aspects of observed aggression. *Journal of Personality and Social Psychology*, 1965, 2, 359-369.
- BROCK, T., & BECKER, L. Debriefing and susceptibility to subsequent experimental manipulations. *Journal of Experimental Social Psychology*, 1966, 2, 314-323.
- FARINA, A., ALLEN, J., & SAUL, B. The role of the stigmatized person in affecting social relationships. *Journal of Personality*, 1968, 36, 169-182.
- FARINA, A., HOLLAND, C., & RING, K. Role of stigma and set in interpersonal interaction. *Journal of Abnormal Psychology*, 1966, 71, 421-428.
- FARINA, A., & RING, K. The influence of perceived mental illness on interpersonal relations. *Journal of Abnormal Psychology*, 1965, 70, 47-51.
- FARINA, A., SHERMAN, M., & ALLEN, J. G. Role of physical abnormalities in interpersonal perception and behavior. *Journal of Abnormal Psychology*, 1968, 73, 590-593.
- FILLENBAUM, S. Prior deception and subsequent experimental performance: The faithful subject. *Journal of Personality and Social Psychology*, 1966, 4, 532-537.
- GOFFMAN, E. *Stigma*. Englewood Cliffs, N. J.: Prentice-Hall, 1963.
- HARTMANN, D. The influence of symbolically modeled instrumental aggression and pain cues on the disinhibition of aggressive behavior. Unpublished doctoral dissertation, Stanford University, 1963. Cited by A. Bandura. Vicarious processes: A case of no-trial learning. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. Vol. 2. New York: Academic Press, 1965.
- MILGRAM, S. Issues in the study of obedience: A reply to Baumrind. *American Psychologist*, 1964, 19, 848-852.
- MILGRAM, S. Some conditions of obedience and disobedience to authority. In I. Steiner & M. Fishbein (Eds.), *Current studies in social psychology*. New York: Holt, Rinehart & Winston, 1965.
- NUNNALLY, J. *Popular conceptions of mental health*. New York: Holt, Rinehart & Winston, 1961.
- NUNNALLY, J., & KITROSS, J. Public attitudes toward mental health professions. *American Psychologist*, 1958, 13, 589-594.
- RING, K., WALLSTON, K., & COREY, M. Mode of debriefing as a factor affecting subjective reaction to a Milgram-type obedience experiment: An ethical inquiry. *Representative Research in Social Psychology*, in press.
- RUBIN, I. Increased self-acceptance: A means of reducing prejudice. *Journal of Personality and Social Psychology*, 1967, 5, 233-238.
- WRIGHT, B. *Physical disability—A psychological approach*. New York: Harper & Row, 1960.
- YABLONSKY, L. Watch out, Whitey. *The New Republic*, 1966, 154, 10-12.

(Received October 4, 1968)

EFFECT OF INFORMATION, SUBJECT AGE, AND INSTRUCTIONAL SET ON PREDICTIVE JUDGMENTS

JAMES L. KNESTRICK¹ AND LEON GORLOW

Pennsylvania State University

The effects of three variables on the accuracy of predictive judgment were examined. The three variables were (a) information, which had three levels—no information given the judges about the person whose behavior was to be predicted, biographical information given them about the person, and information given them about the person which was considered relevant to the behavior to be predicted; (b) subject age, which consisted of two levels—child and adult; and (c) instructional set, which consisted of two levels—no instructions given the judges, and instructions given them oriented toward furnishing a set regarding how to predict. Significant differences were found between the two levels of subject age and also among the three levels of information. The effect of instruction approached significance. There were some additive effects apparent when all three variables were considered at one time.

Research in the area of accuracy of prediction has taken numerous forms. Initially many of the studies dealt with the characteristics of the "good" judge of personality. Intelligence, artistic ability, interest patterns, sex, background, and social attitudes have been found to be related to accuracy of predictive judgments (Cline, 1955; Estes, 1938; Taft, 1956).

More relevant to the present study, however, is some recent research which directs itself not to the characteristics of the judges but toward the *conditions* under which accurate prediction takes place. These studies have examined such conditions as training, information of various types concerning the *S* whose behavior was to be predicted, and feedback concerning accuracy of prediction (Blanchard, 1966a, 1966b; Crow, 1957; Kostlan, 1954; Sechrist, Gallimore, & Hersch, 1967; Sines, 1959). Generally, these studies have indicated that while there has been some progress in determining factors which affect accuracy of interpersonal prediction, the results are equivocal and the methodology questionable. For example, in a study by Fancher (1966), judges' ability to predict increased with feedback, while they did not

in Sechrist et al.'s (1967) study. In addition, for most studies which have been reviewed, little attention has been directed to the important consideration of Bruner and Tagiuri (1954) which involved the selection of appropriate criteria against which to assess the accuracy of prediction.

The present study was designed to determine the effect of three variables on the accuracy with which judges can predict behavior. The three variables under examination were (a) age difference between the *Ss* whose behavior was being predicted, (b) instructional set, and (c) information about *Ss*.

Judges were asked to predict which alternative act an *S* would choose to perform in a series of two-choice situations. The criterion against which to assess the accuracy of prediction was unequivocally determinable by the fact that *S* actually carried out one of the alternatives. This aspect of the study was responsive to the problem of criterion validity—the criterion for assessing the accuracy of prediction was *not* the judgment of an expert.

Age difference between *Ss* was chosen as a variable because research has indicated that the characteristics of the *S* whose behavior is being predicted influence the accuracy of the predictions (Baker & Block, 1957; Estes, 1938; Hamlin, 1954). It was assumed that concomitant with difference in age were other

¹ Now at the Harrisburg State Hospital, Harrisburg, Pennsylvania. Requests for reprints should be sent to James L. Knestrick, 270-E North Arlington Avenue, Harrisburg, Pennsylvania 17109.

characteristics which would influence the accuracy with which the behavior of *Ss* could be predicted. Two age levels were selected, 30 years apart, in order to maximize the effects of this variable.

Instructional set is used in the study to mean giving the judges instructions which might enable them to make more accurate predictions. It was assumed that people will manifest an element of consistency when responding to various related environmental stimuli. This is especially true when the response requires that a performance actually be carried out. Since the behavior which was being predicted in the study required actual performance by the *S* whose behavior was being predicted, it was assumed that his responses would be consistent when responding to similar items. Therefore, instructions were given the judges which alerted them to this assumed consistency. This variable had two levels, one of which was the consistency manipulation just described, and the other was the absence of instructions other than those necessary to describe to the judges what was expected of them.

The third variable under investigation was information about the *S* whose behavior was being predicted. The information took three forms. The first form of information was a control level in which the judges received no information concerning the *S* about whom predictions were to be made. The second was biographical information, consisting of background information such as vocation, family structure, etc. This form of information was used because of research findings indicating that this might increase accuracy of predictive judgment (Kostlan, 1954; Sines, 1959). The third form of information was relevant information, chosen because of its assumed relationship to the behavior which was to be predicted. It consisted of some questions and answers intended to give the judges an idea of the general behavior patterns of the *Ss*.

The overall purpose of the study was to determine the effects of the three variables already discussed on prediction when the behavior to be predicted was behavior which *S* must carry through. This requirement of *performing* the chosen alternative was assumed to influence the choice being made and

to affect accuracy in the judges' predictions when in the presence of the three variables.

METHOD

Two video tape recordings were produced, each of which recorded the behavior of one *S* when he was presented with tasks requiring some physical manipulation. These acts constituted the behavior to be predicted.

The sequence of events in the experimental situation was as follows: Judges entered a room which contained a television monitor and a video tape recorder. Depending on the experimental conditions in effect at the time, one of the two video tapes was played to them, under one of three levels of information and one of two levels of instruction. On the tape the judges saw two people—one was the *E* and the other an *S* to whom 20 two-choice situations were presented. The judges heard and saw the *E* provide the *S* with an opportunity to carry out one of the choices associated with the first task. At that point the tape was stopped and the judges were asked to predict which task *S* would perform. After the judges' predictions were recorded, the tape was started again and the judges were allowed to observe *S*'s choice. Then the second task was presented. This procedure was followed for all 20 two-choice situations.

Video Tape Construction

Two 25-minute, professionally directed video tapes were produced, one with an 11-year-old boy as the *S* and the other with a 43-year-old male adult. The situation which was seen on the tapes by the judges was that of two people facing one another across a table—*E* and *S*. The sequence of events for any one item of the 20 which were presented was: A card appeared on the monitor giving the number of the item to follow, for example, "Item 1"; the card faded out and a brief period of darkness ensued; *E* and *S* appeared sitting at a table facing one another; *E* presented *S* with two choices (this was done both verbally, i.e., "Would you rather do this, or this?" and physically, i.e., *E* held up each choice so that both *S* and the judges could see the materials clearly); the picture faded to black for the judges to record their predictions; the picture reappeared to show *E* asking *S* to choose which alternative to perform; *S*'s actual performance was shown in a close-up shot; the picture faded. This sequence was followed for all 20 items.

The Tasks

Santostefano (1957) investigated forced choice acts as measures of personality. He felt that there was a need for a testing method which

(1) allows for the free expression of personality dimension in situations requiring the subject to cope with people and objects, (2) produces unequivocally identifiable responses, (3) yields mea-

tures the meanings of which are not completely dependent on judgment, and (4) is feasible for use in ordinary clinic and research setting [p. 14].

He empirically determined the personality dimensions represented by choices in 20 two-choice situations by factor analyzing their intercorrelations. As a result of the factor analysis, four relatively well-defined factors emerged. (See Santostefano, 1957, for the factor descriptions and tasks which loaded on each factor.) It was the choices in these 20 two-choice situations, then, which were used in this experiment as the behavior to be predicted.

The Judges

One hundred and twenty male volunteers from introductory psychology classes were used as judges. They were all males in order to control for reported sex differences in the ability to predict behavior (Nidorf & Crockett, 1964; Whalen, 1955). Judges were randomly assigned in groups of 10 to each of 12 experimental conditions. The mean age of the group was 18 years 9 months.

The Variables

Subject age. An 11-year-old boy was selected as one *S* and a 43-year-old male adult as the other. Both were Caucasian. Since similarity between *S* and judge has been found to influence the accuracy of predictive judgment, and since similarity was not under study here, an attempt was made to control for it (Blanchard, 1966a, 1966b; Lundy, 1956; Notcutt & Silva, 1951; Taft, 1966). This was accomplished by determining empirically that adult age which the judges felt was as much unlike them as was an 11-year-old boy. This was done by having 57 *Ss* whose average age was 18 years 7 months respond to a simple questionnaire. They were asked to indicate on a 10-point scale how much they felt like other age levels. The age levels ranged from 5 to 55 years in steps of 5 years. Number 1 on the scale indicated complete similarity and Number 10 complete lack of similarity. The results of this survey showed the mean value for the 10-year age level to be 8.42 and age 45 to be that age level above the age of the responders which they felt was equally unlike themselves as the 10-year level. On this basis the ages 11 and 43 were determined to be equally unlike the judges.

The two *Ss* selected were instructed to respond to the tasks in any way they felt like responding. There was no attempt by *E* to influence their behavior in any way. At the time of the video taping of their responses, they were unaware of the way in which the tape would subsequently be used. This controlled for possible effects this information might have had on their responses. The items were presented in the same order, and the verbal instructions given by *E* were the same in each case.

It should be pointed out that since there was only one *S* to be predicted for each age group, generalizations concerning age as a determinant of ac-

curate predictive judgments are limited. That is, while age is the apparent difference between the two *Ss*, they may differ along some other dimension, unrelated to age, which could also discriminate between people of the same age.

Instructions. The effects of two different types of instructions on the judges were evaluated. Under the assumption that one of the effects of training is to create a "set" that people are relatively consistent in their behavior, the effect of inducing such a set in relatively naive judges was evaluated.

In the first level of instruction, the goal was to inform the judges what their task would be, but to furnish no procedural information which they might use to enhance their prediction. They were given no information concerning the tasks and were given none concerning how they might use the information gained on one task to enhance their prediction on subsequent tasks.

At the second level of instruction, the judges were furnished with some procedural clues and were informed of the purpose of the tasks. The investigators felt that if the judges were aware of the fact that *S's* responses to the tasks might reflect certain aspects of his personality, this might have some effect on the accuracy of their predictions. It was emphasized in these instructions that procedurally, consistency might be the best frame of reference from which to make their predictions.

Information. It is reasonable to assume that the more information we have about a person the more accurate we should be in attempting to anticipate or "predict" his behavior. Accuracy of prediction, however, does not depend on magnitude of information alone. Other aspects of information such as relevance and type also play a role in influencing accuracy of prediction. Blanchard (1966a, 1966b) found that relevant information tended to enhance accuracy of prediction. Both Kostlan (1954) and Sines (1959) found background or biological information useful in the prediction process when compared with other sources of information, for example, tests, etc. With these findings in mind, it was decided to investigate the relative effects of three types of information on accuracy of prediction. The behavior to be predicted was that which has been described above. The three types of information used in this experiment were as follows:

Level I: No information. Under this condition the judges received no information concerning the person whose behavior was to be predicted, other than his age. They were informed of his age for reasons discussed under "Subject age."

Level II: Biographical information. Under this information condition, the judges received information about the *S* which could not be considered directly related to the behavior to be predicted. Judges were given biographical information consisting of age, family description, and professional or vocational data. This information was supplied on a sheet of paper attached to the judges' answer sheets. Before the experiment began, *E* read aloud one of the two, depending on which tape was to be played.

TABLE 1
ACCURACY SCORES UNDER EACH
EXPERIMENTAL CONDITION

Subject	Information		
	No information (c ₁)	Biographical information (c ₂)	Relevant information (c ₃)
No instructions (b ₁)			
Child (a ₁)	11,11,10,10,9 8,8,6,6,6 NC = 85	11,10,10,10,9 9,7,7,7,6 NC = 86	13,10,10,10,10 10,10,9,8,6 NC = 96
Instructions (b ₂)			
Child	10,10,9,9,9 8,8,8,8,7 NC = 86	14,12,10,10,10 9,9,9,8,7 NC = 98	12,12,11,11,10 9,9,9,8,8 NC = 99
No instructions			
Adult (a ₂)	13,12,12,11,11 10,10,8,8,7 NC = 102	17,14,13,12,12 12,12,11,11,8 NC = 122	15,15,15,14,12 12,11,10,9,9 NC = 122
Instructions			
Adult	17,15,14,13,11 10,10,9,7,5 NC = 111	16,15,14,14,13 12,11,11,10,7 NC = 123	15,15,15,14,14 13,13,13,13,12 NC = 137

Note.—NC = number of correct predictions per cell.

The judges were requested to follow along as it was read.

Level III: Relevant information. As will be recalled from the discussion of the tasks, Santostefano (1957) found four discriminating factors to be present in the responses of individuals to the 20 tasks. On the basis of Santostefano's descriptions, the investigators have chosen to label the factors: Factor I—destruction aggression versus nondestruction aggression; Factor II—threat avoidance versus risk taking; Factor III—extroversion versus introversion; Factor IV—covert aggression versus order and structure of environment. The authors felt that by having Ss respond to questions which reflected their tendencies relative to the above factors and by informing the judges of their responses, direct and relevant information concerning the predictive task would be conveyed. The questions for both Ss were the same. These questions and answers were intended to communicate to the judges a picture of the general behavior patterns of the Ss; that is, they were designed to tap certain personality dimensions but were not meant to furnish information which was obviously related to the specific tasks. If these forced choice acts do discriminate between "types" of people in terms of their response tendencies, then Ss responding to them should manifest some degree of consistency in their replies. Therefore, information furnished by S, if it reflects similar response tendencies in other areas of behavior, should enhance a judge's accuracy when he is

predicting that S's responses to these forced choice acts. Two weeks elapsed between the time S responded to the tasks and the time he was asked to respond to the questions. This was done to minimize the effects of one upon the other. The verbatim responses of Ss were used without editing. As in Information Level II, the judges were presented with these questions and answers prior to the experiment. The E read aloud both question and answer and had the judges follow along. They were then instructed to read them again to themselves.

Experimental Design and Summary of Procedure

There were three factors under investigation in the study, with 2, 2, and 3 levels each, respectively. The design was factorial, taking the form of a $2 \times 2 \times 3$ factorial analysis of variance with 10 Ss per cell.

The same procedure was followed for all 12 experimental sessions. Judges entered a room in which 10 chairs had been placed in front of a television monitor. One of the two video tapes (representing the two levels of the subject age variable) was already loaded on an Ampex video tape recorder. Pencils and an answer sheet were distributed to each S. Depending on under which condition of instruction the session was to be conducted, one of the two sets of instructions was stapled to the answer sheet. Also stapled to the answer sheet was one of the levels of information. (If the condition called for Level I of information, nothing was supplied.) The first page of the handout was the instruction page. The E read aloud the appropriate instructions and requested that the judges follow along. After E had assured himself that the instructions were understood, the judges were asked to turn the page to the information section. After one reading aloud by E, the judges were asked to read it again to themselves. After all questions had been answered, the judges were cautioned to do their own work and to refrain from talking during the showing of the video tape. The tape was then started and the first item appeared. After the S on the tape had been presented with the two alternatives and the picture had faded to black, the tape was stopped by E and the judges were asked to indicate on their answer sheets which choice they felt S would make. There was no time limit set, but on the average all judges made their responses within 30 seconds. They could at any time refer to the instruction sheet or the information sheet. The order of presentation of items on the tape was randomly determined.

RESULTS

The measure of accuracy of prediction was the total number of correct predictions made by the judges in each of the 20 situations presented. Table 1 reports these accuracy scores for each experimental condition.

Table 2 provides the summary of the analysis of variance. The comparison between the two levels of subject age (Factor A) was significant. The F ratio resulting from this comparison was 51.19, which is significant at $\alpha .01$ level. This difference was the result of greater accuracy of prediction when the S being observed was the adult S .

The comparison between the two levels of instruction (Factor B) was not significant. The F ratio was 3.09; an F ratio of 3.93 is required to obtain significance at the .05 level. The direction of the obtained difference, however, indicated greater accuracy of prediction under the condition of "instruction" as opposed to the condition of "no instruction."

The effect of different types of information (Factor C) was significant at $\alpha .01$, with an F value of 6.93. Because of the significant F ratio, the Newman-Keuls test was conducted to determine the magnitude of the differences between the treatments taken two at a time. Table 3 shows the results. With the level of significance set at .05, biographical information was found to be significantly effective in increasing accuracy of prediction when compared with the no-information condition. The relevant information condition was the most significant condition when compared with the no-information condition. A comparison between biographical information and relevant information was found not to be significant. However, inspection reveals a greater number of correct predictions under the relevant information condition.

TABLE 2

SUMMARY TABLE FOR ANALYSIS OF VARIANCE

Source	df	SS	MS	F
Subject (A)	1	232.42	232.42	51.19*
Instruction (B)	1	14.02	14.02	3.09 ^a
Information (C)	2	62.93	31.46	6.93*
A \times B	1	.66	.66	
A \times C	2	7.10	3.55	
B \times C	2	.79	.39	
A \times B \times C	2	7.38	3.69	
Within (error)	108	490.30	4.54	
Total	119	815.60		

* 3.93 required for significance at .05.

^a $p < .01$.

TABLE 3

MATRIX OF DIFFERENCES BETWEEN TREATMENTS:
RESULTS OF NEWMAN-KEULS TEST

Ordered treatment totals	A	B	C
No information (A) = 384	—	45*	70*
Biographical information (B) = 429		—	25
Relevant information (C) = 454			—

Note.—Differences necessary for significance at .05 between A and B, A and C, and B and C are 37.85, 45.40, and 37.85, respectively.

* Significant at $\alpha = .05$.

DISCUSSION

Subject Age

It has been suggested by previous research in this area that perhaps the characteristics of the person being judged or about whom predictions are being made may influence the accuracy of the judgments or predictions (Baker & Block, 1957; Hamlin, 1954). As shown in Table 2, very significant differences were found between the two levels of subject age employed in the study. The judges were a great deal more accurate when predicting the behavior of the adult than when predicting the behavior of the child. One possible explanation for these results is that the judges were more accurate when predicting the behavior of the adult because they felt more similar to the adult. However, a control for similarity was employed in the study. It is felt that this difference can be more adequately explained as a function of subject characteristics which are concomitant with age. In an attempt to determine what these characteristics may be, each S 's responses to the 20 items were examined. Certain expectancies exist relative to S 's responses to these items. These expectancies are determined by the clustering of certain of the items on different factors. As mentioned before, four discriminating factors were believed to be reflected in these clusters. For example, on Factor I, five of the two-choice situations were found to cluster together. The clustering was determined by the fact that people who chose Alternative A on one item would also tend to choose a particular alternative on some other item. Thus, when five items were found to cluster or "load" on Factor I,

TABLE 4

SUBJECTS' RESPONSES TO ITEMS LOADED ON FACTOR I

Situation	Responses	
	Child	Adult
1. a. Breaking large light bulb. b. Breaking small light bulb.	b	b
2. a. Breaking light bulb. b. Watering plant.	b	b
3. a. Driving in spike. b. Pulling out spike.	a	b
4. a. Striking sheet of glass. b. Refusing to strike glass.	a	b
5. a. Examining with hand interior of box. b. Looking at concealed picture.	b	b
	3	5

Note.—A person choosing the "a" alternative of the first situation should also choose the "a" alternative of the other situations; similarly for "b."

this meant that there was a tendency for people responding to these items to respond in a "consistent" manner to them. For example, if one of the Ss chose to break the light bulb when faced with the choice of breaking a light bulb or watering a potted plant we would expect him also to choose to drive in a spike when faced with the choice of driving in a spike or pulling out a spike. The responses of the Ss used in this experiment were examined in order to determine whether their responses manifested this element of consistency. Table 4 shows the items that clustered together and each S's responses to those items for Factor I. It is apparent that the adult S was more consistent in his responses to those five situations than the child. If we were to assign a value of 1 to

each consistent response, the adult's score for Factor I would be 5 and the child's score would be 3. Using this method of tabulation on all factors, the results are shown in Table 5.

The investigators feel that these differences in consistency for the most part account for the difference between the two Ss in terms of how accurately their behavior was predicted. It may be that the younger S has not yet established enduring patterns of responding when dealing with environmental stimuli as has the adult. It may also be argued that the type of tasks used in the study do not tap the areas of the child's behavior which are characterized by consistency. If, for example, the two-choice items had had a play versus work quality, the child might have responded consistently and his behavior would then have been more accurately predicted. This suggests that in the prediction of behavior, S's age in combination with the type of behavior being predicted may play a vital role in the accuracy with which predictive judgments are made. These findings support those of Hamlin (1954), Baker and Block (1957), and Estes (1938).

A word of caution is also in order at this point. Because practical considerations limited the number of video tapes which were made, we have only one S for each of the two age groups. Any discussion must take this fact into consideration. A number of replications using different Ss are required before one can say unequivocally that the subject age is that which determines consistency or that it is age which determines the accuracy of predictive judgments.

TABLE 5

CONSISTENCY OF SUBJECT RESPONSES
FOR THE FOUR FACTORS

Factor	Subject		Number of situations
	Adult	Child	
I	5	3	5
II	3	2	4
III	1	1	2
IV	4	3	5

TABLE 6

TOTAL NUMBER OF CORRECT PREDICTIONS
FOR EACH SUBJECT UNDER TWO
LEVELS OF INSTRUCTION

Subject	Level		Overall increase
	No instruction	Instruction	
Child	267	283	16
Adult	346	371	25

Instructions

The rationale for most training programs in clinical psychology, psychiatry, social work, and related disciplines is that somehow a person can be taught the skill which is required to perform his professional function. In all cases, one of these required skills is the ability to make accurate judgments about people. It is assumed that people who receive the appropriate training are better able to make these judgments than those who receive no training. These considerations led to including instructional set as a variable to be assessed in the present experiment.

As shown by the results, the differences between these two levels of instruction failed to reach significance. They were, however, in the direction of more accurate prediction under the "instruction" level, with an F value of 3.09. An inspection of the effect of instruction for each S taken independently indicates that this manipulation may have been more effective when the adult S 's responses were being predicted. This is to be expected, since it was shown in the previous discussion that the adult *was* more consistent than the child. However, when this manipulation is considered independently for the child, we find a trend on the part of the judges to be more accurate in their predictions when receiving instructions.

Considering the facts that the overall differences *approached* significance and that where the S responded in a consistent manner there was greater accuracy, we are hesitant to conclude that instructional set had no effect. It appears more likely that the judges did indeed respond to the instructions and were able to make more accurate predictions under this condition; however, the degree of inconsistency manifested by the child S tended to decrease the overall effects. Since the child was not completely inconsistent in his responses, there was some increase in accuracy in predicting his behavior when instructions were given. Obviously, if there had been a significant effect as a function of the particular S whose behavior was being predicted under a particular level of instruction, there would have been a significant interaction between these two variables. As can

be seen in Table 3, this interaction was not significant.

These results suggest that perhaps the rationale behind training programs such as those mentioned above is valid and the ability to make accurate predictive judgments can be taught or communicated to the novice under the appropriate conditions, that is, correct instructions—instructions which are easily communicated and which are relevant to the behavior to be predicted.

Information

The third variable under investigation was information. The results of the analysis of variance show that the effect of the three levels of information was significant. The differences were in the direction of increased accuracy of prediction with different amounts and kinds of information. Table 3 shows the total number of correct predictions under each level of information and the test of ordered means conducted to determine where the differences occur.

The statistically significant difference between Levels I and II would suggest that accurate background information concerning S tends to enhance accuracy of prediction, whether or not it is related directly to the behavior to be predicted. The writers feel that the information furnished under Level II was used by the judges to create a stereotype and their predictions were based on what they felt a person like that would do when asked to respond to the 20 tasks. These results tend to support those of Kostlan (1954) and Sines (1959).

Level III of information (relevant information) appeared to be the most effective treatment. When compared with Level I, there was a statistically significant increase in accuracy under this condition. It also appears to be superior to Level II, but not at a statistically significant level. Here, the judges apparently used S 's statements about his feelings and his behavior to form a hypothesis concerning how he would respond to the 20 items. These hypotheses were more often correct than not, as evidenced by the results.

The effect of information on accuracy of prediction suggests that information sufficient

to create a stereotype can be effective in enhancing predictive accuracy. Also, it appears that information which is directly related to the behavior to be predicted is the most effective kind of information.

Combined Effects

When all three variables are considered, there appear to be additive effects. For example, the cell or treatment combination in which the smallest number of correct predictions was made was child $S(a_1)$, no instruction (b_1), no information (c_1). The treatment combination in which the greatest number of correct predictions was made was adult $S(a_2)$, instructions (b_2), relevant information (c_3). This result indicates that while the variables have an effect on accuracy when considered alone, there appears to be an "ideal" combination of the three conditions.

The conclusion can therefore be drawn that when behavior is being predicted, the accuracy of the prediction will depend to a great extent on (a) the age of the person whose behavior is to be predicted—it must be considered in light of the kind of behavior which is being predicted; (b) the presence of an accurate instructional set concerning how to predict; and (c) the amount of relevant information available to the predictor about the person whose behavior is to be predicted.

REFERENCES

BAKER, B. O., & BLOCK, J. Accuracy of interpersonal prediction as a function of judge and object characteristics. *Journal of Abnormal and Social Psychology*, 1957, 54, 37-43.

BLANCHARD, W. A. Assimilation and contrast in interpersonal prediction with control for the interaction of real similarity and differential accuracy. *Journal of Personality and Social Psychology*, 1966, 3, 567-573. (a)

BLANCHARD, W. A. Relevance of information and accuracy of interpersonal prediction: A methodological note. *Psychological Reports*, 1966, 18(2), 379-382. (b)

BRUNER, J. S., & TAGIURI, R. The perception of people. In G. Lindzey (Ed.), *Handbook of social*

psychology. Vol. 2. Reading, Mass.: Addison-Wesley, 1954.

CLINE, V. B. Ability to judge personality assessed with a stress interview and sound film technique. *Journal of Abnormal and Social Psychology*, 1955, 50, 183-187.

CROW, W. J. The effect of training upon accuracy and variability in interpersonal perception. *Journal of Abnormal and Social Psychology*, 1957, 55, 355-359.

ESTES, S. Judging personality from expressive behavior. *Journal of Abnormal and Social Psychology*, 1938, 33, 217-236.

FANCHER, R. E. Explicit personality theories and accuracy in person perception. *Journal of Personality*, 1966, 34, 252-261.

HAMLIN, R. M. The clinician as a judge: Implication of a series of studies. *Journal of Consulting Psychology*, 1954, 18, 233-238.

KOSTLAN, A. A. A method for the empirical study of psychodiagnosis. *Journal of Consulting Psychology*, 1954, 18, 83-88.

LUNDY, R. M. Assimulative projection and accuracy of prediction in interpersonal perceptions. *Journal of Abnormal and Social Psychology*, 1956, 52, 33-38.

NIDORF, L. J., & CROCKETT, W. H. Some factors affecting the amount of information sought about others. *Journal of Abnormal and Social Psychology*, 1964, 69, 98-101.

NOTCUTT, B., & SILVA, A. L. M. Knowledge of other people. *Journal of Abnormal and Social Psychology*, 1951, 45, 30-37.

SANTOSTEFANO, S. Forced-choice acts as measures of personality. Unpublished doctoral dissertation, Pennsylvania State University, 1957.

SECHRIST, L., GALLIMORE, R., & HERSCH, P. D. Feedback and accuracy of clinical predictions. *Journal of Consulting Psychology*, 1967, 31, 1-11.

SINES, L. K. The relative contribution of four kinds of data to accuracy in personality assessment. *Journal of Consulting Psychology*, 1959, 23, 483-492.

TART, R. Some characteristics of the good judge of others. *British Journal of Psychology*, 1956, 47, 19-29.

TART, R. Accuracy of empathic judgments of acquaintances and strangers. *Journal of Personality and Social Psychology*, 1966, 3, 600-604.

WHALEN, J. J. An investigation of the influence of certain factors on individual differences in predictive behavior. Unpublished doctoral dissertation, Pennsylvania State University, 1955.

(Received October 4, 1968)

RORSCHACH PREDICTION OF SUCCESS IN CLINICAL TRAINING:

A SECOND LOOK ¹

RAE CARLSON ²

California State College, Fullerton

This study reports the development of a Rorschach Index based on an ego-psychological conceptualization of an "optimal personality picture" of the successful trainee in clinical psychology. Rorschach Index predictions for 155 trainees of the Veterans Administration assessment project were compared with dichotomous predictions from the Miller Analogies Test (MAT) and the Strong Vocational Interest Blank (SVIB). The Rorschach Index predicted success and failure and identified failures for "personality reasons" more effectively than the MAT or the SVIB. Implications for use of an "ideal type" approach in prediction are discussed.

A major and controversial finding emerging from the Clinical Psychology Assessment Project (Kelly & Fiske, 1951) was the failure of projective tests to predict differential success among clinical trainees. The present study reports previously unpublished data on the prediction of progress in clinical training with a Rorschach Index based on ego-psychological constructs.

The scope, methods, and findings of the assessment project (fully described in Kelly & Fiske, 1951) are so well known as to require only a brief summary here. Graduate students admitted to the Veterans Administration (VA) training program in clinical psychology in the late 1940's were studied in a 5-year assessment program with the primary purpose of "the evaluation of a variety of procedures as predictors of later success in graduate training and professional functioning in clinical psychology [p. 193]." A wide variety of objective and projective tests, subjective and quantitative methods were used in

making prediction ratings. While assessment staff members were encouraged to use their preferred personality theories and constructs in formulating evaluations of Ss, they were required to express their evaluations on a standard set of 8-point rating scales. Criterion measures, including objective tests and annual ratings by university and installation supervisors, were also expressed in terms of standard 8-point scales. Results of the study indicated that only intellectual aspects of success in clinical training (as contrasted with clinical and social skills) could be predicted reliably; only the Miller Analogies Test and the Strong Vocational Interest Blank had any predictive value. Rorschach predictions were particularly fallible, showing no significant relationship to any criterion measures.

The present investigation, undertaken as a subproject of the VA assessment project, made two major departures from the orientation of the parent study: (a) in contrast to the eclectic, empirical approach of the larger study, predictions of success or failure were based on an explicit, theoretically derived formulation of an "ideal type"; (b) qualitative patterns, based on dichotomized variables, were used for both prediction and outcome measures, in contrast to the quantitative, dimensional approach of the larger study.

When the prediction situation is seen from the perspective of ego psychology, the immediately relevant assessment problem is that of identifying candidates whose career com-

¹Preparation of this manuscript was aided by a United States Public Health Service Special Fellowship administered by Educational Testing Service, Princeton, New Jersey. Research reported in this study was supported by the Veterans Administration under Contract VAm-22556 with the University of Michigan. The writer is grateful to E. Lowell Kelly and to Ruth Bishop Heiser for their support and encouragement in the conduct of the study, and to William Ward and Roy Shore for their comments on an earlier draft of this paper.

²Requests for reprints should be sent to Rae Carlson, who is now at 4323 Rosedale Avenue, Bethesda, Maryland 20014.

mitment is based on mature ego-identity.³ One looks for candidates who give evidence of (a) general maturity, differentiation, and ego integration, and (b) consonance between personal qualities and the demands of the career training situation so that the professional task will demand and use potentialities, and can become a basis of continuing ego-integration. An "optimal personality picture" of the successful trainee (described below), based on a conceptual analysis of the requirements of clinical training, guided the development of the Rorschach Index. This basic rationale further assumes that (a) evidences of strengths and adaptive resources are more important than evidences of weakness, conflict, or defensiveness, and (b) given the presence of the criterial personality pattern, great diversity in other personality characteristics may be tolerated—while, conversely, an S's possession of many "positive," valued characteristics is irrelevant if the basic "optimal personality picture" is lacking.

The departure from a dimensional treatment of prediction and outcome measures was dictated by the view of selection-prediction questions as intrinsically dichotomous decisions about individuals. One asks: "Is this person a good bet?" "Will he make it or not?" The alternative kind of question asked in the VA assessment study ("How well does a position on Scale A predict a position on Scale B?"), although supported by a more advanced measurement technology, is seen as fundamentally instrumental for making decisions of a quantitative kind.⁴

According to this rationale, a personality measure specifically designed to assess a candidate's "fit" to the hypothetical optimal personality picture should be more clearly related to success in clinical training than measures lacking a conceptual basis in the criterial personality pattern. Thus, the hypothesis to be tested is that the Rorschach Index developed in this investigation (a) pre-

dicts success or failure in clinical training, and (b) discriminates successful and unsuccessful trainees more frequently than the "best predictors" of the VA assessment project—the Miller Analogies Test and the Strong Vocational Interest Blank.

METHOD

Development of the Rorschach Index

The procedure used in the development of the Rorschach Index included the following steps: formulation of an "optimal personality picture" which should characterize the successful trainee, translation of this schema into the language of the Rorschach, and development of an objective Rorschach Index on the basis of empirical study of a small number of cases.

Three sets of assumptions guided the formulation of an optimal personality picture: (a) "Base-rate" assumptions—Since all of the trainees were college graduates who sought professional training in clinical psychology and had been screened and accepted by the Veterans Administration and by participating universities, it was assumed that the group would be adequate—and relatively homogeneous—with respect to intellectual ability and freedom from psychopathology. Therefore, many personality characteristics of clinical importance (e.g., adequacy of reality testing) could be considered peripheral to this prediction problem. (b) Ego integration—Qualities of psychological differentiation and integration considered criterial for mature ego-identity status included evidence of inner resources and stability, mature expression of affect, and effective and flexible controls, with a minimum of constriction or rigidity. (c) Situational requirements—Success in clinical training demands a specific set of personality characteristics. The program is first of all an academic graduate program dealing with "psychological" content at an abstract level; satisfaction and success in such a program should require intellectual ability used in goal-directed striving, and an introversive orientation. The development of effectiveness and responsibility in clinical skills further demands sensitivity and maturity in interpersonal relationships and an adequate and flexible system of controls.

In translating this optimal personality picture into the language of the Rorschach,⁵ emphasis was placed on "structural" characteristics of the psychogram, in contrast to content, and on the organization of psychological processes, reflected in ratios, as opposed to absolute indicators. Traditional indexes of conflict, anxiety, and defensiveness were explicitly deemphasized. The "ideal candidate," then, should present a Rorschach psychogram in which human movement responses (representing inner resources, introversive orientation, the cognitive aspect of goal-directed

³ A succinct definition of Erikson's (1950) concept of ego-identity is offered by Douvan and Adelson (1966, p. 15): "The identity concept, in short, concentrates on the fusion of these elements (identifications, capacities, opportunities, and ideals) into a viable self-definition."

⁴ An eloquent statement of a similar rationale has been presented by Tyler (1959).

⁵ Cf. Klopfer, Ainsworth, Klopfer, and Holt (1954) for scoring and interpretation of Rorschach responses as used in this study.

striving) are high; form-color and form-texture responses (representing mature expression of affect and dependency) predominate over relatively immature color-form and texture-form responses; and human movement and form-color responses (representing psychological differentiation and integrative controls) adequately balance form responses.

Development of an explicit, objective Rorschach Index of this optimal personality picture proceeded as follows: Four case records of trainees (including the Rorschach protocol and other "raw" assessment data, but excluding any interpretations, prediction, or criterion ratings) were examined blindly⁶ by the investigator, who predicted success or failure and attempted a provisional statement of explicit Rorschach criteria on the basis of the pattern described above. Since these blind predictions proved accurate in discriminating the two PhDs and two dismissals in this group, two additional cases were submitted for blind predictions and refinement of Rorschach criteria. Again, these cases were successfully predicted, and a final statement of Rorschach criteria was developed to separate the successful and unsuccessful candidates in the total sample of six trainees. The resulting Rorschach Index consisted of five characteristics (based on Klopfer scoring⁷ of main responses only). All five of the following characteristics must be present in a Rorschach record for a prediction of "success" in clinical training:

$$M \geq \text{sum } C$$

$$M + FC \geq CF + C$$

$$FC \geq CF + C$$

$$F\% < 50\%$$

At least one *FC* and one *Fc* response

Reliability of scoring of the Rorschach Index was established by submitting a random sample of 16 Rorschach records from the total VA sample to an experienced clinician for scoring and prediction on the basis of the Rorschach Index. Agreement between predictions of the independent judge and the investigator was obtained in all cases.

Subjects

The *Ss* included 155 male graduate students enrolled in the VA training program in clinical psychology who had participated in the 1946 and 1947 assessment programs. This sample included all trainees for whom Rorschach protocols and relevant criterion data were available. *Ss*, assessment procedures, and criterion measures are described in Kelly & Fiske, 1951).

⁶ *E* was aware that the first group of four cases included two successful and two unsuccessful trainees; predictions for the final set of two cases were entirely blind.

⁷ Since the Rorschach Index was developed prior to the appearance of Klopfer's revised system (Klopfer et al., 1954), it does not include such current refinements in scoring or interpretation as form-level rating or Klopfer's formulation of ego psychology and Rorschach hypotheses.

Procedure

Rorschach protocols of all trainees were scored in terms of the Rorschach Index and predictions of success or failure made on this basis. In a few cases where original Rorschach records had been scored by Beck or Hertz systems, the investigator rescored records according to the Klopfer system in order to apply the Rorschach Index.

Dichotomous predictions of success or failure were made from the Miller Analogies Test (MAT) and the Strong Vocational Interest Blank, VA Clinical Psychologist Key (SVIB), as follows: A cutting score of 69, empirically determined to give the best prediction, was established for the MAT. SVIB scores had been previously coded on a standard 8-point scale; predictions of success were made for *Ss* receiving a coded score of five or above.

Predictions of success or failure in clinical training were evaluated for two criterion groups: Group I (PhD vs. dismissal) included the 68 trainees who had either received the doctorate in clinical psychology or had been dismissed from the program by the spring of 1950. This group included trainees who had entered the program at various levels and who had participated in the 1946 or 1947 assessment programs. Group II (successful vs. unsuccessful) consisted of 105 first-year trainees from the 128 *Ss* assessed in the summer of 1947. The *Ss* were eliminated if they had previous experience with the Rorschach ($n = 8$) or relevant criterion data were incomplete ($n = 15$). The *Ss* were classified as successful if they had received the doctorate in clinical psychology or received a 1950 median university rating of five or above on the three component skills of diagnosis, therapy, and research. Unsuccessful *Ss* were defined as those who had resigned or had been dismissed from the program or had received a 1950 median university rating of less than five on competence in diagnosis, therapy, and research.

Predictions of success or failure were evaluated with chi-square or binomial tests. Since all hypotheses involved directional predictions, one-tailed tests were used.

RESULTS AND DISCUSSION

Table 1 summarizes Rorschach Index and MAT predictions for Group I. The data indicate that the Rorschach Index predicts success in clinical training, identifying early PhD's as well as the MAT, and identifying early failures significantly better than the MAT (binomial probability $< .02$).

Is the Rorschach Index sensitive to relevant personality characteristics "missed" by the MAT? One source of evidence on this point is available in the analysis of the 34 dismissals. According to the basic rationale, the Rorschach Index should be relatively more effective in identifying candidates whose

TABLE 1
CORRECT PREDICTIONS OF SUCCESS OR FAILURE FOR GROUP I

Predictor	N	PhD predicted correctly		N	Dismissal predicted correctly		N	Total predicted correctly		χ^2
		N	%		N	%		N	%	
Rorschach										
Index	34	28	82	34	28	82	68	56	82	29.9**
MAT	33	27	82	34	21	62	67	48	72	10.1*

* $p < .01$.
** $p < .001$.

failures had been categorized by universities as due to "personality reasons" in contrast to "academic reasons." Predictions from the Rorschach Index and the MAT for the two types of dismissals are summarized in Table 2.

While the majority of dismissals of each type were correctly identified by both the Rorschach Index and the MAT, only Rorschach Index predictions of "personality" failures attained statistical reliability.

Results for the more homogeneous sample of first-year trainees in Group II are summarized in Table 3. Since MAT and SVIB scores were not available for all Ss, comparisons are based on percentages of correct predictions of success or failure. Table 3 indicates that the Rorschach Index gave a significantly higher proportion of correct predictions of failure than the SVIB (binomial probability .005) or the MAT ($p < .005$), while the SVIB and the MAT were more accurate in predicting success (binomial probabilities of $p < .08$ and $p < .06$, respectively).

One of the most striking aspects of the results is the sharp reduction in the accuracy of prediction from either the Rorschach Index

or the MAT as one moves from Group I to Group II (binomial probabilities are .007 and .02, respectively), despite the fact that the two criterion groups contain some of the same individuals. While we have argued for the clear-cut PhD-dismissal criterion (which in this context means *early* PhD vs. *early* dismissal) as the more relevant measure of success or failure, it is also true that Group I represents greater heterogeneity in terms of age, experience, and ability. Therefore, the issues of heterogeneity-of-sample and conceptual-relevance-of-criteria need to be examined separately. Fortunately, it is possible to do so by comparing results of the present investigation with previously reported findings from the VA assessment project on Rorschach Index predictions against three sets of criterion data for the same homogeneous sample of trainees represented in Group II of this study. Relevant comparisons are summarized in Table 4.

Table 4 indicates clearly that the Rorschach Index predictions for the same Ss are more accurate as the stringency and differentiation of criterion measures increase. "Professional Promise," the criterion most successfully predicted by the Rorschach Index, is based on a higher cutting score and uses the rater's judgment of relevant weighting of performance on component skills for the individual case. "Successful versus Unsuccessful" (Group II of the present study) is based on a cutting score of five—the midpoint of the scale—and employs an arbitrary equal weighting of three component skills. "Pass-Fail," as a criterion, assumes that trainees who have left the program (for whatever

TABLE 2
PREDICTIONS OF FAILURE FOR SUBJECTS
LATER DISMISSED

Predictor	N	"Academic reasons" predicted correctly	$p^a <$	N	"Personality reasons" predicted correctly	$p^a <$
Rorschach Index	14	64	.22	20	90	.001
MAT	13	61	.30	20	65	.14

^a Binomial test.
^b Includes "personality" and "combination of personality and academic reasons."

TABLE 3
CORRECT PREDICTIONS OF SUCCESS OR FAILURE FOR GROUP II

Predictor	N	Successful predicted correctly		N	Unsuccessful predicted correctly		N	Total predicted correctly		χ^2
		N	%		N	%		N	%	
Rorschach										
Index	40	23	58	65	46	71	105	69	66	8.24**
MAT	40	27	68	57	31	54	97	58	59	4.54*
SVIB	40	24	60	63	34	54	103	58	56	2.46

* $p < .05$.** $p < .01$.

reason) are "failures" and that candidates still in training after 3 years (regardless of their level of competence) are likely to receive the doctorate, and are, therefore, "successful."

Clearly, the Rorschach Index is less generous in predicting success than the MAT or the SVIB; predictive errors tend to be in the direction of "false negatives." This is not surprising, of course, since the Rorschach Index was designed to identify only the most promising candidates. However, base rates for prediction of success vary with the maturity and heterogeneity of the sample. Where success was predicted for 50% of the trainees in Group I, the base rate is significantly reduced to 40% among the younger, more homogeneous trainees of Group II (binomial probability .05). Predictive accuracy similarly decreases (from 82% to 66%; binomial probability .007) when the maturity-heterogeneity of the sample and stringency of criteria are decreased simultaneously.

These considerations suggest that the Rorschach Index demands a degree of maturity, differentiation, and experience-in-living which may not be readily found in beginning graduate students. (It should be noted that even the first-year trainees of Group II included many veterans with some years of wartime experience in addition to undergraduate training.) While the apparent sensitivity of the Rorschach Index to something like "maturity of character development" may be counted as evidence of its construct validity, this consideration could limit its usefulness in the selection of beginning graduate students.

Obviously the major strength of the

present Rorschach Index—its derivation from a custom-made conceptualization of the demands of clinical training—becomes a barrier to its application in other prediction situations. One could not expect this measure to serve as a general assessment technique, nor to identify promising candidates in other professional fields with different patterns of personality requirements. Moreover, since the rationale of the present approach requires the conceptualization of intrinsic situational demands prior to development of predictive instruments, it is quite possible that this Rorschach Index may no longer predict success in current clinical training programs. The striking shifts in the *Zeitgeist* and technology of clinical psychology over the past 20 years may well have altered the pattern of demands and satisfactions to be found in clinical training and the characteristics of those who seek such training.

However, the general approach developed in this investigation may hold promise for a variety of assessment and prediction problems. The basic approach—that of formulating an "optimal personality picture" as a

TABLE 4
RORSCHACH INDEX PREDICTIONS FOR GROUP II
ON THREE CRITERION MEASURES

Criterion	Basis for prediction of success: PhD or	ϕ
Professional promise*	Rated overall clinical competence ≥ 6	.39**
Successful- Unsuccessful	Median rating on three component skills ≥ 5	.28*
Pass-Fail*	Still in training in third year	.13

* For details, see Kelly and Fiske (1951, pp. 191-192).

* $p < .01$.** $p < .001$.

basis for prediction—is independent of the ego-psychological constructs or the Rorschach test used here. The present findings, however, offer encouragement for the use of ego-psychological constructs—even at this relatively superficial level—as a source of relevant questions and for the Rorschach as an instrument for answering these questions.

REFERENCES

- DOUVAN, E., & ADELSON, J. B. *The adolescent experience*. New York: Wiley, 1966.
- ERIKSON, E. *Childhood and society*. New York: Norton, 1950.
- KELLY, E. L., & FISKE, D. W. *The prediction of performance in clinical psychology*. Ann Arbor: University of Michigan Press, 1951.
- KLOPPER, B., AINSWORTH, M., KLOPPER, W. G., & HOLT, R. *Developments in the Rorschach technique*. Yonkers-on-Hudson, New York: World Book, 1954.
- TYLER, L. Toward a workable psychology of individuality. *American Psychologist*, 1959, 14, 75-81.

(Received October 4, 1968)

VERBAL AND OVERT-BEHAVIORAL ASSESSMENT OF A SPECIFIC FEAR¹

ANTHONY F. FAZIO²

University of Wisconsin, Milwaukee

Using college students ($N = 70$), self-ratings of fear on a modified Fear Survey Schedule and overt approach-avoidance responses to a large insect were investigated. Professed fears of insects were negatively related ($p < .001$) to approach responses toward them. Sex of the *E* and whether or not *E* handled the insect were not associated with significant differences in approach behavior. Behavior which might be described as at least mildly phobic occurred in 27.1% of the *Ss*. Implications of the diagnostic utility of the questionnaire and the behavioral test were discussed.

Directly or indirectly the principal criterion used to evaluate the efficacy of psychotherapy has been the verbal report (e.g., the patient claims he is or is not feeling better). The utility of behavior modification techniques can be assessed, however, directly from objective changes in overt behavior (Ullmann & Krasner, 1965, p. 28). Phobic behaviors are readily interpreted in behavioral terms as avoidance responses made in the presence of identifiable stimuli. Typically, in controlled investigations of behavior modification, patients or *Ss* are interviewed and confronted with an anxiety producing situation in order to assess (e.g., Lazarus, 1961) or verify (e.g., Hogan & Kirchner, 1967) their professed fears. Nevertheless, *Ss* accepted for interviewing are often preselected on the basis of self-evaluative measures. Even though questionnaires and self-rating scales produce quantitative data quickly and easily, their relationships to actual overt behavior require evaluation.³ Investigating public speaking fears, Paul (1966, p. 129) reported a significant correlation ($r = .41$, $p < .001$) between a self-report measure of confidence and behav-

ioral checklist ratings by judges. Lang (1968), however, has commented on the generally nonsignificant correlations between fear-survey questionnaires and overt avoidance behaviors (e.g., for fear of snakes, $r = .26$, $p < .10$).

The first objective of the present study was to evaluate the relationship between professed or denied fears of a specific stimulus object (cockroaches) and overt approach-avoidance responses toward that object. Questionnaire items dealing with stimuli similar to and different from the phobic object were also examined.

Of the many situational characteristics which may effect *S's* behavior, two were examined in the present study: sex of the *E* and whether or not *E* handled the insect. Bandura, Grusec, and Menlove (1967) have reported avoidance responses to be significantly modified by the modeling behavior of a peer. If behavioral tests are to be used as diagnostic tools, the extent to which therapeutic treatments are confounded with testing-situation influences needs to be assessed.

The third objective was to obtain normative data on avoidance responses of college students toward a relatively large insect.

METHOD

Fifty-six female and 14 male introductory psychology students voluntarily participated in the "interview and survey." The insect, a harmless variety of cockroach, about 2½ inches long and 1 inch wide (*Gromphadorhina portentosa*), was kept in a 14 × 8½ × 10 inch empty glass aquarium with a wire screen cover. The container was placed on a table behind a partition.

The Fear Survey Schedule (FSS-III; Wolpe & Lang, 1964) was used as a vehicle for obtaining

¹ This research was supported by the graduate school of the University of Wisconsin, Milwaukee. The author wishes to thank Allan Wicker and Thomas Stampfl for their comments on this manuscript. The assistance of Kathlyn Schroeder and Thomas Cotey is gratefully acknowledged.

² Requests for reprints should be sent to Anthony F. Fazio, Department of Psychology, University of Wisconsin, Milwaukee, Wisconsin 53201.

³ Although Lanyon and Manosevitz (1966) obtained both a self-report measure of fear and a behavioral measure, they did not report the degree of correlation between the sets of scores.

TABLE 1

PRODUCT-MOMENT CORRELATIONS BETWEEN
QUESTIONNAIRE ITEM OR ITEM COMPOSITES
AND RATED OVERT TEST BEHAVIOR

Items	<i>r</i>
1. Cockroaches	-.37*
2. Large ugly insects	-.47**
3. Harmless spiders	-.46**
4. Worms	-.16
5. Flying insects	-.49**
6. Crawling insects	-.49**
Composite insect score	-.41**
7. Speaking in public	+.19
8. Entering a room where other people are already seated	-.16
9. People in authority	+.19
10. Being watched working	+.19
11. Looking foolish	+.22
Composite social situation score	+.16

Note.—*N* = 70.

**p* < .002.

***p* < .001.

verbal responses about the particular stimulus object. The schedule contains 72 items (e.g., cats, being alone, blood) to be rated from "Not At All" to "Very Much" in 5 steps, on the degree to which each stimulus situation provoked fear or was experienced as unpleasant. The higher the score, the greater the degree of fear inferred. Three additional items were added to the survey: Cockroaches, Large ugly insects, and Harmless spiders.

For the overt-behavioral test, a 7-step rating scale was devised which allowed *E* to identify *S*'s responses relatively objectively:

1. Refused to move to a position from which the insect could be seen.
2. Refused to stand close enough to the table to be able to touch the container.
3. Refused to remove the screen cover.
4. Tip of fingers failed to pass upper rim of container.
5. Fingers did not touch surface or appendages of insect.
6. Failed to grasp insect in firm hold long enough for insect to be picked up.
7. Insect lifted so that all parts were not touching any surface of container.

Procedure

The male and the female *E* assigned *Ss* in order of appearance to experimental conditions in which the *E* handled (*H*) or did not handle (\bar{H}) the insect. The *E* sat adjacent to *S* at a secretarial desk from which the insect could neither be seen nor heard.

For both *H* and \bar{H} *Ss*, *E* said:

This survey concerns fears and/or unpleasant feelings experienced by college students, caused by

things, places, animals, etc. Extensive research of such fears is presently being conducted at a variety of universities. Such fears are neither uncommon among college students nor shameful. To begin with, we would like you to fill out this standardized questionnaire. When you have completed it, we will continue.

They then completed the modified FSS-III. Not scoring their responses, *E* continued:

In an attempt to assess the extent to which students might be afraid of handling certain animals, we have selected one situation listed in the survey you have just completed. We are offering all students the opportunity to handle a completely harmless, but large ugly insect. You are completely free to refuse to do so, if you wish. You will in no way be forced to do anything against your wishes. In a glass container on a table to the right of this partition is a relatively large insect, which occasionally makes a hissing sound, but which nevertheless is completely harmless.

In an effort to control the degree to which *Ss* would use *E*'s behavior as a cue, the instructions explicitly informed the *H Ss* that *E* was going to handle the insect, and informed the \bar{H} *Ss* that *E* was not allowed to handle it. Except for these statements the instructions were as similar as possible.

For *H Ss*, *E* said:

First, we would like you simply to look at it; second, touch it, after I have touched it. Finally, pick it up off the base of the container, after I have picked it up. You are, of course, free to refuse any of these requests.

For \bar{H} *Ss*, *E* said:

First, we would like you simply to look at it; second, touch it. Finally pick it up off the base of the container. You are, of course, free to refuse any of these requests.

For *H Ss*, *E* said:

I will touch the insect, then you may touch the insect. It really cannot hurt you. I will pick the insect up, then you may do so. It is quite harmless.

For \bar{H} *Ss*, *E* said:

Because my behavior may influence yours, I am not allowed to touch the insect. It really cannot hurt you. You may pick up the insect now. It is quite harmless.

The *E* recorded all of *S*'s comments and scored *S*'s approach-avoidance behavior. The test was ended when *S* picked up the insect, or *S* refused to continue, or after 5 minutes had elapsed.

RESULTS

Questionnaire scores for separate items and specific thematic combinations of them were correlated with rated overt test behavior for all 70 Ss (Table 1). The inserted items and scores for two additional items dealing with buglike objects were found to be significantly related to test behavior. These items were selected a priori on the basis of their similarity to the test object. In each case, the more S claimed to fear the insect, the less he or she approached and/or handled it. Verbal measures of fear were found to possess discriminant validity (Campbell & Fiske, 1959) by comparing the correlation for the composite of the insect items with a correlation for a composite of noninsect items. Five items dealing with a theme of feeling conspicuous before a group of observing persons were selected prior to examining the data. The theoretically irrelevant items dealing with social situations and, thus, their composite were not significantly correlated with the specific behavioral test involving a large roach (Table 1).

The predictive utility of selecting Ss for interview on the basis of different degrees of professed fear was also evaluated. If only those indicating "Very Much" fear of, for example, "Crawling insects," were selected, 10% of the total sample ($N = 70$) would have been interviewed. However, this subsample contained only 15.8% of those in the total sample who behaved like phobics—a poor yield. If Ss indicating at least "A fair amount" of fear had been selected, 50% of the sample would have been interviewed. However, this procedure would have netted 89.5% of the Ss rated phobic-like on the basis of the behavioral test.

Except for one male S reporting "Much" fear of "Harmless spiders," and one reporting "A fair amount" for "Large ugly insects," none of the other male Ss indicated more than "A little" fear of any of the insect items, even though two male Ss refused to place their hands in the container. These findings are consistent with the report of Manosevitz and Lanyon (1965) concerning females expressing significantly greater anxiety in general than males.

TABLE 2

DISTRIBUTION OF OVERT-BEHAVIORAL RATINGS

Scale steps						
1	2	3	4	5	6	7
Would not look at insect	Not touch container	Refused to remove screen cover	Fingers not pass upper rim	Not touch surface insect	Not pick up insect	Lifted insect
$N = 1$	$N = 1$	$N = 5$	$N = 12$	$N = 1$	$N = 20$	$N = 30$

Note.—Total $N = 70$.

The data on situational effects were analyzed only for female Ss because there were few male Ss ($N = 14$) and because their verbal and overt-behavioral scores reflected significantly less anxiety ($p < .01$) and avoidance ($p < .05$) than the scores of female Ss. A 2×2 analysis of variance revealed no significant differences in mean rated test behavior associated with sex of E ($F = 1.51$, $df = 1/52$, $p > .10$), handling of the insect ($F = .70$, $df = 1/52$, $p > .20$), or interaction between them ($F = 1.56$, $df = 1/52$, $p > .10$). For all 56 female Ss, the mean and standard deviation of the overt-behavioral ratings were 5.54 and 1.55, respectively.

In spite of the gruesome appearance of the insect, 71.4% of the Ss touched it when asked to, 42.8% picked it up, and only 2.8% refused to even stand near the container. From the distribution of behavioral scores (Table 2) it might be the case that the behavior of placing one's hand in the container divides Ss into groups possibly described as phobic versus nonphobic. However, this dichotomy may also have been an artifact of the difficulty in scoring this particular behavior (i.e., E had to judge if S's hand came close to but did not touch the insect).

DISCUSSION

Although verbal reports to relevant items correlated significantly with overt behaviors, at best only 25% of the behavioral variance was accounted for by this measure. Wicker (1970), in a review of research on attitude-behavior consistency, reports that attitude-behavior correlation coefficients rarely exceed .30, and often are near 0.0. He suggests that

if the stimuli in the verbal situation were more similar to those in the overt response situation, these correlations could be larger. For example, Ss could be asked to describe what they would do *if* they were asked to handle the insect, being tested on an independent occasion.

A second approach toward improving the magnitude of these correlations is suggested by the work of Endler and his colleagues (Endler & Hunt, 1966; Endler, Hunt, & Rosenstein, 1962). In general, they have obtained support for the claim that self-report inventories containing a variety of stimulus situations and response alternatives are associated with more valid personality descriptions. More specific self-descriptions, in turn, might allow for attitude-behavior correlations of larger magnitudes.

One aspect of the present investigation concerned possible reactive error effects associated with *E* (cf. Webb, Campbell, Schwartz, & Sechrest, 1966, pp. 12-34). At the beginning of the interview, Ss were told they were participating in a survey concerning their fears, and were asked to describe (i.e., rate) them. After completing the modified FSS-III, they were informed that a test was arranged to verify at least one of their claims. This procedure can be interpreted to have produced substantial reactive effects, similar to response sets, in the direction of behavioral consistency. Thus, while the correlations obtained here were statistically significant, their relative magnitudes may decrease when other procedures are used.

The specific situational variations manipulated had no significant effect on S's behavior, while individual differences in professed fear of insects accounted for a greater proportion of the behavioral differences. These findings suggest that the behavioral test might be more sensitive to differences in individuals and previous treatments than it is to extraneous characteristics of the situation, thus qualifying as a diagnostic device. However, modeling effects of *E* handling the insect cannot be ruled out when repeated tests are employed. In a review of research on imitation, Flanders (1968) has classified modeling practices such as handling the phobic object as instances of vicarious reinforcement, and

Bandura et al. (1967) have presented data suggesting that repeated exposures to a positive model significantly affect avoidance responses of children toward dogs. College students were used in the present study and were exposed only once to the modeling influences of *E*. Geer and Turteltaub (1967), using female undergraduates and fear of snakes, examined the modeling effects of a confederate who behaved "quite frightened" in one experimental condition, and displayed "very little fear" in another condition. High-fear Ss observing the low-fear confederate were reported to have moved closer to the snake on a postobservation test than high-fear Ss observing the high-fear confederate. All low-fear Ss, defined by touching the snake on the first behavior test, touched it on the second. Although behavioral tests are generally intended to assess rather than modify behavior, investigations using nontreated but repeatedly tested control groups will require checks on the effects of modeling.

The present sample did not contain a sufficient number of male Ss to definitively test possible interaction effects between sex of *E* and sex of *S*. One important characteristic of research involving phobic-like behaviors in nonclinical populations is the tendency for male Ss to perform, or at least attempt, almost every task requested. Efforts to increase the range of behavioral measures (i.e., remove the ceiling effect) have involved items of behavior which require *S* to handle the object for longer time periods and/or in some particularly vulnerable manner (e.g., next to one's face) (cf. Bandura, 1968).

Of the Ss tested, 27.1% behaved in a manner which might be described as mildly phobic (i.e., they refused to place their hand in the container). Lang and Lazovik (1963) report that about 3% of a college sample are to some extent snake phobic. Whereas clinical descriptions (e.g., phobic) do not allow data to be compared objectively, operational definitions involving overt behaviors are important precisely because they allow such comparisons to be made both within and between investigations. Thus the data in Table 2 may be compared with data on, for example, fear of rats (e.g., Hogan & Kirchner, 1967, p. 109).

REFERENCES

- BANDURA, A. Modelling approaches to the modification of phobic disorders. In R. Porter (Ed.), *Ciba foundation symposium: The role of learning in psychotherapy*. Boston: Little, Brown, 1968.
- BANDURA, A., GRUSEC, J., & MENLOVE, F. L. Vicarious extinction of avoidance behavior. *Journal of Personality and Social Psychology*, 1967, 5, 16-23.
- CAMPBELL, D. T., & FISKE, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, 56, 81-105.
- ENDLER, N. S., & HUNT, J. McV. Sources of behavioral variance as measured by the S-R inventory of anxiousness. *Psychological Bulletin*, 1966, 65, 336-346.
- ENDLER, N. S., HUNT, J. McV., & ROSENSTEIN, A. J. An S-R inventory of anxiousness. *Psychological Monographs*, 1962, 76(17, Whole No. 536).
- FLANDERS, J. P. A review of research on imitative behavior. *Psychological Bulletin*, 1968, 69, 316-337.
- GEER, J. H., & TURTELTAUB, A. Fear reduction following observation of a model. *Journal of Personality and Social Psychology*, 1967, 6, 327-331.
- HOGAN, R. A., & KIRCHNER, J. H. Preliminary report of the extinction of learned fears via short-term implosive therapy. *Journal of Abnormal Psychology*, 1967, 72, 106-109.
- LANG, P. J. Fear reduction and fear behavior: Problems in treating a construct. *Research in Psychotherapy*, 1968, 3, 90-102.
- LANG, P. J., & LAZOVIX, A. D. Experimental desensitization of a phobia. *Journal of Abnormal and Social Psychology*, 1963, 66, 519-525.
- LANYON, R. I., & MANOSEVITZ, M. Validity of self-reported fear. *Behaviour Research and Therapy*, 1966, 4, 17-24.
- LAZARUS, A. A. Group therapy of phobic disorders by systematic desensitization. *Journal of Abnormal and Social Psychology*, 1961, 63, 504-510.
- MANOSEVITZ, M., & LANYON, R. I. Fear survey schedule: A normative study. *Psychological Reports*, 1965, 17, 699-703.
- PAUL, G. L. *Insight versus desensitisation in psychotherapy: An experiment in anxiety reduction*. Stanford: Stanford University Press, 1966.
- ULLMANN, L. P., & KRASNER, L. *Case studies in behavior modification*. New York: Holt, Rinehart & Winston, 1965.
- WEBB, E. J., CAMPBELL, D. T., SCHWARTZ, R. D., & SECHREST, L. *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago: Rand McNally, 1966.
- WICKER, A. W. Attitudes versus actions: The relationships of verbal and overt behavioral responses to attitudinal objects. *Journal of Social Issues*, 1970, in press.
- WOLFE, J., & LANG, P. J. A fear survey schedule for use in behavior therapy. *Behaviour Research and Therapy*, 1964, 2, 27-30.

(Received October 4, 1968)

RELATIONS BETWEEN PERSONALITY AND SMOKING BEHAVIOR IN PREADULT SUBJECTS¹

GENE M. SMITH²

*Department of Anesthesia, Massachusetts General Hospital
Harvard Medical School*

The present study of 562 high school and junior high school students yielded results which cross-validate and extend the generality of earlier work with 1,462 adults. In both studies, smoking status was assigned on the basis of self-report information, and personality scores were derived from peer ratings. With the preadults, as with the adults, smokers scored significantly ($p < .001$) lower on measures of "Agreeableness," and "Strength of Character," and scored significantly higher on measures of "Extraversion." In addition, the smokers (in both studies) scored significantly higher than the nonsmokers on the variables "crude," "happy-go-lucky," and "frank." Among preadults, multiple discriminant analyses permitted smoking status to be assigned with accuracy ranging from 65% to 79%. Although most information concerning the psychodynamics of smoking has been obtained from studies of adults, the present study supports the use of such information in developing antismoking educational campaigns aimed at preadults.

In an earlier study of 1,462 college and nursing students, Smith (1967a) found highly significant differences between smokers and nonsmokers regarding personality characteristics, measured both with self-report questionnaires and with the peer rating method. Multiple discriminant analyses performed allowing five measures of personality, derived from peer ratings, to enter the prediction equation permitted smokers to be discriminated from nonsmokers with 68% accuracy, and permitted heavy smokers (20 or more cigarettes a day) to be discriminated from nonsmokers with 76% accuracy. This is distinctly greater accuracy than that obtained in any earlier study. In terms of sample separation, an accuracy level of 68% is obtained when the means for the two samples are approximately 1.0 standard deviations apart. Similarly, 76% accuracy is obtained when the

two samples are approximately 1.5 standard deviations apart. The relationship between sample separation and accuracy of classification is discussed in detail elsewhere (Smith, 1967a).

The present investigation tests the generality of the earlier adult results by replicating the study in a sample of 562 preadults. The literature contains 31 primary reports of empirical studies of the relations between personality and smoking behavior,³ but only two of these studies used preadult Ss (Salber & Rochman, 1964; Stewart & Livson, 1966). Thus, the present investigation of 562 high school and junior high school students adds to this small literature.

The questions of whether, and how, smokers differ from nonsmokers regarding personality is of psychological interest for two reasons. First, the association between smoking and disease, although now widely accepted, is not uniformly interpreted. For example, Berkson (1958) and Fisher (1959), both statisticians, suggest that such genotypic factors as constitution and temperament might predispose some persons to become smokers and also predispose them to develop lung cancer. Whether or not that position has merit, the differential morbidity rates of smokers and

¹ This work was supported in part by Grant TH-3 from the American Cancer Society, in part by Public Health Service Grant GM-15904, and in part by a grant from the Council for Tobacco Research-U.S.A. William G. Cochran, Professor of Statistics, Harvard University, and Frederick Mosteller, Professor of Mathematical Statistics, Harvard University, served as consultants and their advice is gratefully acknowledged.

² Requests for reprints should be sent to Gene M. Smith, Department of Anesthesia, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts 02114.

³ The 31 references may be obtained by writing the author.

nonsmokers will probably become more understandable as constitutional and psychosocial differences between these groups are further specified. Second, little is currently known about the motivational factors underlying decisions to become a smoker, or not; to continue, or quit; to reduce, or not. This is of considerable practical importance because the effectiveness of antismoking educational campaigns will probably improve monotonically with the accumulation of dependable and comprehensive information concerning the psychodynamics of smoking.

METHOD

Measurement of Smoking Behavior

Students who reported that they currently smoke seven or more cigarettes per week were classified as smokers; those reporting less frequent current smoking (or none) were classified as nonsmokers. With this definition (Definition 1), 27% of the high school students and 16% of the junior high school students were classified as smokers. Because frequency of smoking among the junior high school students was low by this definition, the nonsmokers were further divided into those who reported that they had smoked 20 or more cigarettes in their lifetime and those who reported that they had not. Definition 2, which included as smokers those who reported having smoked 20 or more cigarettes in their lifetime, resulted in 38% of the junior high school students being classified as smokers. For junior high school students, smoker-nonsmoker comparisons were made in terms of both definitions. Better discrimination was obtained with Definition 2 than with Definition 1. With Definition 2, 42 of the 84 biserial correlation coefficients (between personality and smoking status) were significant at or beyond the .05 level. With Definition 1, 34 of the 84 coefficients were significant. The junior high school data reported in Tables 1 and 2 are based on Definition 2.

Measurement of Personality

Each student was asked to examine each of 42 bipolar personality traits (Cattell, 1957) and identify, for each trait, the five members of his class who are most like the left pole and the five who are most like its opposite on the right. (Class size ranged from 20 to 35). Selections on the left were considered positive nominations and those on the right were considered negative. The positive and negative nominations a ratee received were scored +1 and -1, respectively. Failure to be nominated was scored zero. Thus, in a peer group of 25, a ratee's score on any given trait could range from +24 to -24. Further details concerning the method of deriving personality scores from peer ratings, the

safeguards employed to reduce bias, and the theoretical rationale underlying the use of peer ratings are given by Smith (1967b).

RESULTS

A factor analysis (principal components rotated to an oblique solution using the bi-quartimin criterion) of the entire sample of peer rating data (Grades 8-12) revealed four factors which agree well with results obtained in earlier studies of nursing students and college students (Smith, 1967b). The relations between smoking and personality are shown in Table 1 for each of the 42 peer rating variables separately; but to facilitate summarization, those variables are grouped by factor. The variables loading positively on each factor are listed separately from those loading negatively.

The values shown in Table 1 are biserial correlation coefficients. Smokers were coded +1 and nonsmokers were coded 0; hence, plus correlation coefficients indicate that smokers have more of the trait than nonsmokers and negative coefficients indicate the opposite.

The results in Table 1 are mutually supportive within factors and highly consistent across the four samples. Each of the six variables loading positively on Factor I (Agreeableness) correlates negatively with smoking status in each of the four samples. These correlations indicate that smokers score lower on "Agreeableness" than do nonsmokers. Each of the five variables loading negatively on Factor I correlates positively with smoking status in each sample. These correlations indicate that smokers score higher than nonsmokers on "Disagreeableness." Stated differently, 44 correlations in the table relate smoking status to "Agreeableness" and all 44 indicate that smokers score lower on "Agreeableness" than do nonsmokers. Of those 44 coefficients, 38 are significant at or beyond the .05 level.

The results for Factor II (Extraversion) are also internally consistent over variables and over samples. Of the 23 (out of a possible 40) correlations which reach the .05 level of significance, 22 indicate that smokers score higher on "Extraversion" than do nonsmokers.

TABLE 1

BISERIAL CORRELATIONS BETWEEN PERSONALITY TRAITS AND SMOKING STATUS
(SMOKER VS. NONSMOKER) IN FOUR SAMPLES OF STUDENTS

Factor	High school students		Junior high students	
	Females (N = 198)	Males (N = 188)	Females (N = 96)	Males (N = 80)
I: Agreeableness				
Positive variables				
Adaptable	-.20**	-.18*	-.35***	-.09
Tender	-.40***	-.24***	-.44***	-.34**
Self-effacing	-.23***	-.20**	-.40***	-.22*
Trusting	-.24***	-.24***	-.31**	-.21
Good-natured	-.24***	-.19**	-.26**	-.10
Considerate	-.36***	-.24***	-.37***	-.22*
Negative variables				
Jealous	+.22**	+.18*	+.28**	+.13
Assertive	+.22**	+.27***	+.27**	+.11
Attention-seeking	+.28***	+.23***	+.35***	+.27*
Demanding	+.26***	+.18*	+.28**	+.27*
Hides mistakes	+.31***	+.25***	+.35***	+.12
II: Extraversion				
Positive variables				
Gregarious	+.14*	+.23***	+.18	+.12
Gay	+.23***	+.16*	+.31**	+.17
Talkative	+.23***	+.15*	+.30**	+.26*
Conventional	-.25***	-.09	.00	-.05
Happy	+.10	+.08	+.16	+.21
Interest in opposite sex	+.11	+.23***	+.33***	+.19
Negative variables				
Reserved	+.11	-.03	+.06	-.09
Quiet	-.29***	-.23***	-.47***	-.19
Shy	-.31***	-.25***	-.40***	-.25*
Thoughtful, pensive	-.17*	-.19**	-.25*	-.22*
III: Strength of Character				
Positive variables				
Conscientious	-.39***	-.33***	-.45***	-.32**
Self-reliant	-.17*	-.17*	-.29**	+.05
Responsible	-.37***	-.24***	-.43***	-.13
Orderly	-.33***	-.22**	-.40***	-.13
Socially mature	-.22**	-.09	-.23*	-.05
Resourceful	-.19**	-.14*	-.31**	-.11
Inquisitive	-.01	+.01	+.15	+.08
Obedient	-.35***	-.23***	-.41***	-.20
Negative variables				
Quitting	+.33***	+.22**	+.38***	+.08
Languid	+.13	+.12	+.11	-.04
Obstructive	+.26***	+.15*	+.22*	+.01
Crude	+.36***	+.24***	+.34***	+.14
Prone to daydream	+.20**	+.16*	+.25*	-.05
IV: Emotionality				
Positive variables				
Emotional	+.12	+.10	+.37***	+.20
Tense	+.18*	+.08	+.12	-.13
Easily upset	-.13	-.17*	+.15	-.17
Negative variable				
Tolerant of stress	-.08	-.10	-.25*	+.02
Complex variables				
Happy-go-lucky	+.32***	+.22**	+.41***	+.19
Frank	+.15*	+.12	+.33***	+.13
Imaginative	+.20**	-.03	+.20*	-.09
Fastidious	-.08	+.07	+.11	-.03

Note.—Plus signs (+) indicate higher scores for smokers than for nonsmokers; minus signs (-) indicate the opposite.
* $p = .05$, ** $p = .01$, *** $p = .001$.

For Factor III (Strength of Character), 33 of the 52 correlations were significant and all 33 indicate that smokers score lower on "Strength of Character" than do nonsmokers.

The results for Factor IV (Emotionality) are suggestive but inconclusive. Only four of the 16 correlations are significant. Three of those indicate higher emotionality among smokers than among nonsmokers, but one indicates the opposite.

Table 1 also shows (see bottom) results for four complex variables (happy-go-lucky, frank, imaginative, and fastidious) which load heavily on more than one of the four factors. Where significant, the correlations involving these complex variables indicate that relative to nonsmokers, smokers score higher on "happy-go-lucky," "frank," and "imaginative."

The main difference between the factor analytic results of the present study and the earlier results based on adults is that the earlier fifth factor, "Refinement," did not appear in the present analysis. One of its two component variables, "fastidious," was found here to be a complex variable which loaded on three of the four factors identified; the other, "crude," was found to load on the factor called "Strength of Character." In the present study, as in the earlier one, smokers scored significantly higher than nonsmokers on the variable "crude" (see Table 2).

Multiple discriminant analyses, using the step-up procedure,⁴ were performed to determine the degree to which smokers could be distinguished from nonsmokers on the basis of personality scores. To avoid capitalizing on what might be called "idiosyncratic, sample-bound relations" between personality and smoking status, a prediction equation was developed empirically, using half of the high school sample ($N = 193$), and this equation

TABLE 2

MEDIAN BISERIAL CORRELATIONS BETWEEN
PERSONALITY TRAITS AND SMOKING STATUS
(SMOKER VS NONSMOKER) IN FIVE
ADULT AND FOUR PREADULT
SAMPLES

Factor	Adults ($N = 1029$)	Preadults ($N = 562$)
I: Agreeableness		
Positive variables		
Adaptable	-.10***	-.19***
Tender	-.17***	-.37***
Self-effacing	-.25***	-.23***
Trusting	-.27***	-.24***
Good-natured	-.13***	-.22***
Considerate	-.29***	-.30***
Negative variables		
Jealous	+.19***	+.20***
Assertive	+.18***	+.25***
Attention-seeking	+.28***	+.28***
Demanding	+.25***	+.27***
Hides mistakes	+.15***	+.28***
II: Extraversion		
Positive variables		
Gregarious	+.21***	+.16***
Gay	+.24***	+.20***
Talkative	+.17***	+.25***
Conventional	-.13***	-.07
Happy	+.10***	+.13***
Interest in opposite sex	+.29***	+.21***
Negative variables		
Reserved	-.02	+.02
Quiet	-.23***	-.26***
Shy	-.29***	-.28***
Thoughtful, pensive	-.21***	-.21***
III: Strength of Character		
Positive variables		
Conscientious	-.30***	-.36***
Self-reliant	-.08**	-.17***
Responsible	-.27***	-.31***
Orderly	-.31***	-.28***
Socially mature	+.01	-.16***
Resourceful	-.02	-.17***
Inquisitive	+.07*	+.05
Obedient	-.33***	-.29***
Negative variables		
Quitting	+.27***	+.28***
Languid	-.02	+.12**
Obstructive	+.14***	+.10***
Crude	+.18***	+.21***
Prono to daydream	+.06*	+.18***
IV: Emotionality		
Positive variables		
Emotional	+.12***	+.16***
Tense	+.06*	+.10*
Easily upset	-.11***	-.15***
Negative variable		
Tolerant of stress	-.09**	-.09*
Complex variables		
Happy-go-lucky	+.20***	+.27***
Frank	+.18***	+.14***
Imaginative	+.01	+.09*
Fastidious	+.06*	+.02

Note.—Plus signs (+) indicate higher scores for smokers than for nonsmokers; minus signs (-) indicate the opposite. In the study of 1462 adults, 433 contributed self-report personality data but no peer rating data.

* $p = .05$.
** $p = .01$.
*** $p = .001$.

⁴ The first variable to enter the prediction equation was the one (out of 42) which had the highest correlation with the criterion. The next to enter was the one having the highest correlation with the criterion after the effect of the first predictor variable was partialled out. At the next and at each succeeding step the entering variable was the one having the highest correlation with the criterion when the influence of the multivariate prediction battery obtained at the preceding step of the analysis was partialled out (Rao, 1952).

was then tested in two cross-validation samples—the second half of the high school sample ($N = 193$) and the sample of 176 students from junior high school. With five of the 42 peer variables in the equation, accuracy of prediction of smoking status was 77% in the first high school sample, 65% in the second, and 67% in the junior high school

sample. When the two cross-validation samples were combined, accuracy of classification of smoking status was 66%. In these analyses males and females were combined to keep the sample sizes relatively large. In earlier analyses, where males and females were analyzed separately, the accuracy of discrimination for high school females was 71% and that for high school males was 69%. The corresponding figures for females and males in the junior high school samples were 79% and 68%, respectively. (The analyses separating the sexes were not cross-validated.) In both univariate and multivariate results, the association between personality and smoking status was greater for junior high school females than for junior high school males.⁵

The data of the present study of preadults were collected with procedures similar to those used earlier in a study of adults. Table 2 reports the results of the two studies. Two biserial correlation coefficients are given for each of the 42 personality variables. The first member of each pair is the median coefficient for five samples of adults studied earlier (Smith, 1967a). The second is the median of the four coefficients obtained in the present study of preadults. The results of the two studies are strikingly similar. Of the 42 pairs of correlations, 39 have the same signs and 34 of these are statistically significant in both studies. Furthermore, the magnitude of the correlation between smoking status and score on a particular personality variable in one study is highly predictive of the magnitude of the correlation for that variable in the other study.

COMMENTS AND CONCLUSIONS

Study of the relations between personality and smoking status in 562 high school and

⁵ Multivariate discrimination for junior high school males was definitely poorer than that for junior high school females—but less so than one might have expected from examining the univariate results. In this regard it should be noted that the smoker/nonsmoker split for junior high school males was 38/42, whereas that for females was 28/68. In addition, the intercorrelation among the 42 peer variables was higher for females than for males. Both of these differences could favor multivariate discrimination for the males. Because of these and other complexities, multivariate results do sometimes depart from expectations based on univariate results.

junior high school students yielded results which are strikingly similar to those found earlier with adults. In both studies the smokers scored higher than the nonsmokers on various measures of "Extraversion," and scored lower on various measures of "Agreeableness," and "Strength of Character." In addition, smokers (in both studies) scored higher than nonsmokers on the variables "crude," "happy-go-lucky," and "frank." Thus the present results, obtained with preadults, cross-validate and extend the generality of the earlier results, obtained with adults.

The results of both studies were (a) highly significant, statistically, (b) internally consistent within factors, and (c) consistent across samples. The uniformly high level of statistical significance shown in Table 2 (Point a just mentioned) is due in part to the large sample sizes. Although a correlation of .11 accounts for only 1% of the criterion variance, it is, nevertheless, significant at the .001 level with a sample of 1,029. The consistency within factors (Point b) is not surprising, since variables comprising each factor are highly intercorrelated; however, the extraordinary consistency across samples (Point c) was surprising; and because of this, we checked the analytic procedures more extensively than usual, but found no evidence of error.

Multiple discriminant analyses permitted smokers and nonsmokers to be discriminated with accuracy ranging from 68% for junior high school males, to 79% for junior high school females. Cross-validation analyses yielded 66% accuracy of classification of smoking status. In the present study of preadults, as in the earlier study of adults, use of peer rating measures of personality, in association with multivariate analyses, provided distinctly greater discriminative power than that reported in any previous study of the relationship between personality and smoking status.

If one believes that the effectiveness of antismoking educational campaigns will increase as more is learned concerning the psychodynamics of smoking, then it becomes important to determine whether information now available from studies of adults can be used to guide campaigns directed at preadults

—a population which has been used only rarely in studies of the psychodynamics of smoking. While the present investigations have dealt with only one aspect of this complex matter, the extraordinary similarity between the results for adults and preadults does provide some justification for using information obtained in studies of adults to guide development of antismoking educational campaigns directed at younger people.

REFERENCES

- BERKSON, J. Smoking and lung cancer: Some observations on two recent reports. *Journal of the American Statistical Association*, 1958, 53, 28-38.
- CATTELL, R. B. *Personality and motivation structure and measurement*. Yonkers, N. Y.: World Book, 1957.
- FISHER, R. A. *Smoking—the cancer controversy*. Edinburgh: Oliver & Boyd, 1959.
- RAO, C. R. *Advanced statistical methods in biometric research*. New York: Wiley, 1952.
- SALBER, E. J., & ROCHMAN, J. Personality differences between smokers and nonsmokers. *Archives of Environmental Health*, 1964, 8, 459-465.
- SMITH, G. M. Personality correlates of cigarette smoking in students of college age. *Annals of the New York Academy of Sciences*, 1967, 142, 308-321. (a)
- SMITH, G. M. Usefulness of peer ratings of personality in educational research. *Educational and Psychological Measurement*, 1967, 27, 967-984. (b)
- STEWART, L., & LIVSON, N. Smoking and rebelliousness: A longitudinal study from childhood to maturity. *Journal of Consulting Psychology*, 1966, 30, 225-229.

(Received October 8, 1968)

EFFECTS OF CUES AND EXAMINER INFLUENCE ON TWO WECHSLER SUBTESTS¹

JEROME M. SATTLER²

San Diego State College

Two experiments studied the effects of cues on the Wechsler Intelligence Scale for Children and the Wechsler-Bellevue Scale. In the first experiment 170 adolescent *Ss* received the Block Design subtest, and results showed that one cue per item did not significantly affect performance on either the initial or repeated administration. In the second experiment 146 adolescent *Ss* received the Block Design and Picture Arrangement subtests, and results showed that a series of cues per item significantly raised scores on both the initial and repeated administrations of both subtests. Test-form order significantly affected scores, while subtest order and *Es* did not.

The administration, scoring, and interpretation of intelligence tests are affected by many variables such as test content, procedures for administration and scoring, *E* competence, and the interpersonal relationship between the *E* and the *S* (Sattler & Theye, 1967). The present investigation studied the effect of a number of these variables on test performance. The primary objective was to determine how systematic alterations in test procedures affect performance on two Wechsler subtests, Picture Arrangement (PA) and Block Design (BD). These two subtests were selected because they especially lend themselves to the development of a series of specific, graduated cues which can be systematically administered. The issue of cue effects arises in part from recent suggestions made by Allison, Blatt, and Zimet (1968) that systematic cues be provided on the BD as part of the standard administrative procedure. Additional objectives were to evaluate the relationship between the Wechsler Intelligence Scale for Children (WISC) and Wechsler-Bellevue (W-B) Scale Form I for the PA and BD subtests, to examine *E* differences, and to study the effects of test form and subtest order on performance.

The following hypotheses were advanced:

¹ This study was supported by Research Projects OE 5-10-429 and OEG-4-7-078057-0402 from the United States Department of Health, Education, and Welfare, Office of Education. The writer is grateful to Edward A. Jacobson and to Leonard L. Tozier for their helpful suggestions.

² Requests for reprints should be sent to Jerome M. Sattler, who is now at Miami University, Department of Psychology, Oxford, Ohio 45056.

(a) Violation of standard procedures by administering cues during the first administration of the subtest(s) has no effect on test performance during either the first administration or a subsequent administration of a comparable form of the subtest(s) given immediately after the first subtest(s). (b) There is no relationship between cues and test performance. (c) The scores obtained by different *Es* do not differ significantly. Two experiments are reported which evaluated the hypotheses. In the first experiment only the BD subtest was used, while in the second experiment the BD and PA subtests were used.

EXPERIMENT I

Method

Subjects. The *Ss* were eighth- and ninth-grade students attending a junior high school in Southern California. The experiment was designed to evaluate the performance of students with average intellectual ability who had no severe emotional or learning difficulties. To satisfy these two criteria, *Ss* were selected if they had School and College Ability Test (SCAT) scores within .60 of a standard deviation from the mean of their class, if they were not attending classes for the educationally or culturally handicapped, or if they did not have a learning disability. Also eliminated from the original sample were students who were foreign born or who had been previously administered the WISC or W-B. Fifty-four *Ss* initially selected had a perfect performance, and they were subsequently eliminated from the sample because the effects of cues could not be evaluated for them. The sample for the statistical analyses consisted of 71 male and 99 female adolescents. Their mean years of age was 14.29 (*SD* = .65 year).

Examiners. Three graduate students, two male and one female, attending San Diego State College

served as the *Es*. The three *Es* were psychology majors who had completed graduate work in intelligence testing and who were working toward their master's degree. When the experiment was initiated only one *E* served in the study, but it soon became evident that additional *Es* would be needed. After the first *E* had tested approximately one-quarter of the *Ss*, the other two *Es* were employed. Since the first *E* participated during the entire study, he tested several more *Ss* than the other two *Es*.

Recording forms with the names of the *Ss* were arranged in a random order, and the *Es* obtained the names of *Ss* by taking approximately five or six record forms for each day's testing from the top of a pile containing all of the record forms. While this procedure was not designed to ensure random assignment of the *Ss* to the *Es*, it eliminated any selection of *Ss* by the *Es*. Each *E* tested both experimental and control *Ss*. *E*₁ tested 20 control and 68 experimental *Ss*, *E*₂ tested 6 control and 33 experimental *Ss*, and *E*₃ tested 7 control and 36 experimental *Ss*.

Procedure. Five groups were used to evaluate the effects of BD cues: four experimental groups each received a different cue, and a fifth served as the control group. Each cue consisted of constructing a specific part of the design in front of *S*. The four cues administered to the respective groups were as follows: showing *S* how to arrange the (a) first row, (b) last row, (c) first column, and (d) last column. Cues were administered immediately following failure of the design.

The *Ss* were assigned separately by sex to one of the five groups. Because the W-B does not take age into consideration when raw scores are converted to standard scores, the sample was divided into 3-month age intervals corresponding to the 3-month intervals on the WISC conversion tables. Using a table of random permutations of nine, each *S* within a 3-month age interval group was randomly assigned to one of the five groups. The order of presentation of WISC and W-B was counter-balanced by assigning odd-numbered *Ss* to the W-B-WISC order and even-numbered *Ss* to the WISC-W-B order.

Each *S* received two administrations of the subtest with one subtest immediately following the other. During the first administration, the experimental groups received the cues, while no cues were administered to the control group. During the second administration, the alternate form of the test was administered, and no cues were presented to any of the groups.

Wechsler's (1944, 1949) discontinuance criteria were used during both the first and second administrations. An item was considered failed when *S* could not complete it correctly on the first administration. Passing the item after the cue was administered was still considered a failure for discontinuance purposes. The time limits stated in the manual were used after the cue was administered.

The *Ss* were tested individually in rooms provided

by the school. The *E* introduced himself to *S* by saying:

Hello, I am Mr. [Miss] _____. I am going to give you a short test in order to compare your present results with previous test scores. Your score today, however, will not be recorded on your school records and will not affect your grades in any way. We would like you to do the best you can. Do you have any questions?

After the testing was completed, *E* said:

Thank you for cooperating. I'd like to remind you that your score does not affect your grades, and it isn't reported on your school records. We are interested only in finding out more information about our tests. We are testing many students at the school, and you were selected by a method something like having your name pulled out of a hat.

The following instructions were read to an *S* when a cue was administered:

I'm going to put together some of the blocks. I will make the bottom row [or top row, left column, right column, depending on the group to which *S* was assigned]. Now you go ahead and finish it. Look at the picture and make one just like it. Tell me when you are finished.

Results^a

Cues. The effectiveness of the four BD cues was evaluated by a three-factor analysis of variance design with repeated measures on one factor. The two independent factors were group and sex of *S*, and administration was the repeated factor. A summary of the analysis of variance appears in Table 1. The only significant factor was the administration: *Ss* achieved significantly higher scores on the second administration ($\bar{X} = 10.81$) than on the first administration ($\bar{X} = 9.26$). Thus, the four cues did not differ among themselves or from the control group in affecting *Ss'* performance on either the initial or second administration, and sex differences were not present.

Examiner variable. The *E* variable was evaluated in two ways, and both used analysis of variance designs. Nonsignificant differences were found in the scores the *Es* obtained and in the number of cues they administered (\bar{X} number of cues = 1.85). The number of cues administered in the two test-form orders and

^a All analyses of variance in both Experiments I and II used the unweighted means analysis procedure because of unequal numbers.

TABLE 1

ANALYSIS OF VARIANCE OF BD SCORES FOR
GROUPS, SEX, AND ADMINISTRATIONS
IN EXPERIMENT I

Source	<i>df</i>	<i>MS</i>	<i>F</i>
Between Ss			
Groups (A)	4	11.00	1.04
Sex (B)	1	29.04	2.75
A × B	4	8.90	.84
Error (between)	160	10.56	1.00
Within Ss			
Administrations (C)	1	199.90	63.55*
A × C	4	1.89	.60
B × C	1	1.54	.49
A × B × C	4	2.99	.95
Error (within)	160	3.14	1.00

**p* < .001.

in the four experimental groups were also found by analysis of variance procedures to be not significantly different.

Order, age, ability, and test-form variables. Analyses of variance were performed to determine whether the order of administering the test forms affected scores and whether the ages and ability level of the Ss differed in the five groups. The results in each case were not significant for these variables. The randomization procedures were thus effective in producing similar ages and ability levels in the groups. A Scheffé test, used to evaluate whether the WISC and W-B total means differed, was not significant ($t = .70$, $df = 160$, $p > .05$); therefore, the two test forms produced similar scores.

EXPERIMENT II

Experiment II was designed to measure the effects of a series of cues rather than only one cue, and to evaluate two subtests, BD and PA. Because the first column cue was found to be the most effective in Experiment I (although not significantly), it was used as a basis for establishing the BD cues in Experiment II.

Method

Subjects. The Ss in Experiment II were similar, with some slight exceptions, to those participating in Experiment I. Students were attending two junior high schools and were in the seventh, eighth, and ninth grades. The procedures described in Experiment I for selecting average ability Ss with

minimal emotional and learning difficulties were also used in Experiment II. None of the Ss had participated in Experiment I. There were 49 Ss, initially selected, who had perfect performance on either the BD or PA, and they were subsequently eliminated from the sample because for these Ss, the effects of cues could not be evaluated. The sample for the statistical analyses consisted of 71 male and 75 female adolescents. Their mean years of age was 13.25 ($SD = .67$ year).

Examiners. Four graduate students, three male and one female, attending San Diego State College, served as the Es. Three of the four Es participated in Experiment I. The fourth E's experience was similar to that of the other three. The first three Es began testing Ss at the same time, and about halfway through Experiment II, the fourth E was employed. The Es obtained the names of their Ss in exactly the same manner as described in Experiment I. E1 tested 17 control and 23 experimental Ss, E2 tested 12 control and 12 experimental Ss, E3 tested 18 control and 23 experimental Ss, and E4 tested 22 control and 19 experimental Ss.

Procedure. The Ss received two administrations of both the BD and PA subtests. During the first administration, the experimental group received cues, while the control group was administered the two subtests in the standard manner. During the second administration, which immediately followed the first, the alternate form of the Wechsler test was administered, and cues were not administered to either of the groups. Four counterbalanced orders were used to administer the subtests: (a) BD-WISC, PA-WISC, BD-W-B, PA-W-B; (b) PA-W-B, BD-W-B, PA-WISC, BD-WISC; (c) BD-W-B, PA-W-B, BD-WISC, PA-WISC; (d) PA-WISC, BD-WISC, PA-W-B, BD-W-B.

When the experimental Ss failed to complete items correctly, they received a series of specific, graduated cues designed to facilitate successful completion of the items immediately after failing an item. Three cues were designed for each subtest. The BD cues were as follows: *Cue 1*—The S received 50% additional time beyond the completion time stated in the manual. *Cue 2*—The E arranged the first column of blocks. *Cue 3*—The E arranged the first column plus an additional number of blocks. The number of blocks arranged in Cue 3 was as follows: three of the four blocks for a 4-block design (first column plus first block of second column), first two columns for a 9-block design, and first three columns for a 16-block design.

The PA cues were as follows: *Cue 1*—The S received 50% additional time beyond the completion time stated in the manual. *Cue 2*—The E arranged a specified number of cards. When the series contained three or four cards, E arranged the first card. When the series contained five cards, E arranged the first and second cards, and when the series contained six cards, E arranged the first, second, and third cards. *Cue 3*—The E arranged a greater number of cards than in Cue 2, with the exception of the series containing three cards. For a three-card series, the number of cards arranged by E was the

same as in Cue 2, because if two cards were arranged in a three-card series, only one card would remain for *S* to arrange. When the series contained four cards, *E* arranged the first and second cards. When the series contained five cards, *E* arranged the first, second, and third cards. When the series contained six cards, *E* arranged the first, second, third, and fourth cards. A specific order for the remaining cards was used based on the order stated in the manual unless the order would have automatically resulted in a correct solution. In these cases a different order was used for the remaining cards.

If *S* successfully completed the item after the first cue, *E* presented the next item. If *S* was unable to complete the item after the first cue, the second cue was administered. If *S* was still unable to complete the item correctly, the third cue was administered. As in Experiment I, Wechsler's (1944, 1949) discontinuance criteria and time limits (except for the first cue) were used during both the first and second administrations. Testing conditions were similar to those described in Experiment I.

Results

Cues. The subtest scores were evaluated by a four-factor analysis of variance design using group and sex as the two independent factors, and administration and subtest as the two repeated factors. The results of the analysis of variance, presented in Table 2, show that three significant *F*s were found which indicated the following: (a) the experimental group ($\bar{X} = 10.10$) achieved significantly higher overall scores than the control group ($\bar{X} = 9.30$), thus the cues improved performance; (b) second administration scores ($\bar{X} = 10.43$) were higher than first administration scores ($\bar{X} = 9.00$); and (c) BD scores ($\bar{X} = 10.04$) were higher than PA scores ($\bar{X} = 9.40$). Because none of the interactions was significant, the results also indicate that the experimental group achieved significantly higher scores than the control group on both the first and second administrations, and that both the experimental and control groups obtained significantly higher scores on the second administration of both subtests.

Examiner variable. The *E* variable was evaluated in three ways, and all used analysis of variance designs. Nonsignificant differences were found among *E*s in the scores they obtained on both subtests, and in the number of cues ($\bar{X} = 3.11$) they administered in the various orders of administration. However, when the number of items on which cues were

TABLE 2
ANALYSIS OF VARIANCE OF BD AND PA SCORES
FOR GROUPS, SEX, SUBTESTS, AND ADMIN-
ISTRATIONS IN EXPERIMENT II

Source	df	MS	F
Between Ss			
Groups (A)	1	92.03	6.40*
Sex (B)	1	17.31	1.20
A × B	1	22.25	1.54
Error (between)	142	14.37	
Within Ss			
Subtests (C)	1	53.83	7.86**
A × C	1	11.48	1.67
B × C	1	24.98	3.64
A × B × C	1	2.32	.34
C × Ss	142	6.84	
Administrations (D)	1	284.26	62.01***
A × D	1	3.44	.75
B × D	1	.04	.00
A × B × D	1	.13	.02
D × Ss	142	4.58	
C × D	1	7.83	1.92
A × C × D	1	.03	.00
B × C × D	1	2.34	.57
A × B × C × D	1	1.96	.48
C × D × Ss	142	4.08	

* $p < .05$.

** $p < .01$.

*** $p < .001$.

administered was used as the dependent variable, three significant *F*s appeared: the *E* factor ($F = 2.98$, $df = 3/52$, $p < .05$), order factor ($F = 3.82$, $df = 3/52$, $p < .05$); WISC BD first $\bar{X} = 2.06$, W-B PA first $\bar{X} = 1.70$, W-B BD first $\bar{X} = 2.39$, WISC PA first $\bar{X} = 1.85$), and the *E* × Order interaction ($F = 2.14$, $df = 9/52$, $p < .05$). The four *E* means within each order were tested by use of the Newman-Keuls procedure. The results indicated that there were no significant differences among the four *E*s' means when the WISC BD or W-B PA was administered first. However, when the W-B BD was administered first, *E*₄'s mean ($\bar{X} = 3.50$) was significantly higher than the means of each of the three other *E*s. The means for *E*₁ ($\bar{X} = 2.08$), *E*₂ ($\bar{X} = 2.00$), and *E*₃ ($\bar{X} = 2.00$) were not significantly different from one another. When the WISC PA was administered first, the mean of *E*₁ ($\bar{X} = 1.00$) was significantly lower than the means of the other three *E*s ($\bar{X} = 2.00$, $\bar{X} = 2.16$, and $\bar{X} = 2.22$, respectively). The results indicated that one of

the four *Es* gave cues on more items than the other three *Es* in one order, and another *E* gave cues on fewer items than the other three *Es* in another order.

Order Variable. The effect of order of administration was evaluated by a three-factor analysis of variance design. The group by order of administration was the independent factor (eight levels), and administration and subtest were the repeated factors. The group by order factor was not significant ($F = 1.30$, $df = 7/138$, $p > .05$), while the administration factor ($F = 8.22$, $df = 1/138$, $p < .01$), subtest factor ($F = 80.91$, $df = 1/138$, $p < .001$), and the Group by Order \times Subtest interaction ($F = 6.50$, $df = 7/138$, $p < .001$) were significant.

The significant interaction indicates that significant differences between the BD and PA subtests are dependent on the administration order. Individual mean comparisons, using the total means for each subtest (first plus second administrations), made between the BD and PA subtests in each of the eight group by orders of administration revealed that significantly higher BD than PA scores were obtained in both the control and experimental groups on the two orders having the W-B first, and that the experimental group obtained significantly higher BD than PA scores on one order which had the WISC administered first. The individual mean comparison results suggest that in the control group there is a greater disparity between the BD and PA subtest scores when the W-B is administered during the initial administration and followed by the WISC than when the WISC is first administered and followed by the W-B. In the experimental group, a significant difference between BD and PA was not related to test-form order, per se.

Test forms Two Scheffé tests were conducted to evaluate whether the subtest means were similar for each form. The average of the four WISC means (first and second administration plus two orders) was compared separately for each subtest to the average of the four W-B means. Neither of the Scheffé tests were significant (for BD, $t = .67$, $df = 138$, $p > .05$, for PA, $t = 11$, $df = 138$, $p > .05$). Thus the two test forms produced means which were not significantly different.

Age and ability variables. Analyses of variance were performed to determine whether the ages and ability level of the *Ss* differed in the groups. The results in each case were not significant.

DISCUSSION

Performance on either the initial or subsequent administration of the BD was not affected in Experiment I when *Ss* were given single cues consisting of constructing part of the design (first or last columns or first and last rows) on a mean number of 1.85 designs. However, when more extensive BD cues were given in Experiment II by providing cues on a mean number of 2.00 designs and providing a mean number of 3.11 cues (i.e., giving more time, constructing first column, and constructing first and part or all of the second and/or third columns), *Ss* scored higher on the first and second administrations of the BD than those *Ss* not receiving cues on the first administration. In Experiment II, PA performance was also significantly improved on both the first and second administrations when a series of cues was administered. These results have implications for the examination procedure.

The results indicate that alterations in standard procedures, when minor, are not likely to affect BD performance; however, when extensive, BD performance is affected. The *E* should therefore not introduce extensive testing-the-limit procedures during the examination; rather, he should wait until the examination has been completed before such procedures are attempted. Allison et al.'s (1968) suggestion that graded help be introduced during the administration of the BD subtest is clearly untenable because the present results show that help procedures can significantly raise performance. The suggestion made by Wechsler (1958) that stories be obtained on the Wechsler Adult Intelligence Scale PA items immediately after each item is completed must also be questioned. This procedure introduces an additional variable which may affect *S's* performance. By telling a story after he has completed his arrangement, *S* has time to evaluate further his arrangement. He may see that his arrangement

was incorrect, and thereby possibly obtain insight which may be useful to him in solving the next arrangements. The results of the present study do not provide an answer to whether test scores are altered when stories are asked. However, this latter modification of test procedure should be investigated for its possible effects on test performance.

Because the data were collected on two subtests from average ability students with minimal learning and emotional difficulties, it is difficult to generalize the findings to other subtests or groups. Other subtests and groups therefore should be studied in future research concerned with the effects of cues on test performance.

Other results point to the following: First, the *E* variable does not necessarily affect the reliability of the intelligence test score, and the sex of the *E* is not necessarily a significant variable either. Second, WISC and W-B BD and PA subtests produced mean scores which were not significantly different. Third, the subtest order of administration and the test-form order of administration are very complex variables which at times may affect *S*'s or *E*'s performance. Departures from the standard subtest order of administration must therefore be carefully evaluated for their possible effects on performance. Much remains

to be learned about how and why different administrative orders affect test performance.

Approximately 25% of those tested were eliminated from the analyses because they passed all of the items on either the BD or PA subtests. Thus, these subtests did not provide adequate ceilings with a population of average ability adolescent students. It is recommended that future studies using the WISC or W-B which focus on the effects of modifying standard procedures use either younger age groups or develop additional items to expand the range of difficulty of the subtests. The ceilings of other tests or subtests may also be limited, and it is recommended that tests used in similar investigations be evaluated for their range of difficulty.

REFERENCES

- ALLISON, J., BLATT, S. J., & ZIMET, C. N. *The interpretation of psychological tests*. New York: Harper & Row, 1968.
- SATTLER, J. M., & THEYE, F. Procedural, situational, and interpersonal variables in individual intelligence testing. *Psychological Bulletin*, 1967, **68**, 347-360.
- WECHSLER, D. *The measurement of adult intelligence*. (3rd ed.) Baltimore: Williams & Wilkins, 1944.
- WECHSLER, D. *Wechsler Intelligence Scale for Children*. New York: Psychological Corporation, 1949.
- WECHSLER, D. *The measurement and appraisal of adult intelligence*. (4th ed.) Baltimore: Williams & Wilkins, 1958.

(Received October 10, 1968)

PERSONALITY CORRELATES OF ACCURATE EMPATHY IN A COLLEGE COMPANION PROGRAM¹

GEORGE J. VESPRANI²

University of Cincinnati

A cross-validation of research by Bergin and Solomon on personality test correlates of empathy was attempted using 33 untrained female college students as quasi-therapists. Their interactions with hospital patients were recorded, sampled, and rated using the Truax Accurate Empathy Scale. Accurate empathy (AE) ratings were correlated with student personality measures from the Edwards Personal Preference Schedule and the Minnesota Multiphasic Personality Inventory (MMPI). The results partially support data reported by Bergin and Solomon in that Depression and Psychasthenia scales of the MMPI were found to be negatively and significantly correlated with AE ratings of students. In addition, this study provides data about changes in levels of AE over time and methodological pitfalls in sampling this variable.

In recent research on those variables that presumably affect the course of psychotherapy, empathy has been shown to be an important and relevant one. Truax (1963), as well as other researchers, has shown that when high levels of empathy are communicated to patients, the likelihood of patient improvement is increased. In fact, Truax (1963) and Bergin (1963) have shown that a low level of empathy on the part of the therapist may actually lead to patient deterioration. It seems imperative, therefore, that all therapeutic endeavors, in an attempt to maximize patient progress, provide a high level of this therapist variable. Certainly, this has relevance for all of the established and accepted forms of therapeutic encounters from usual forms of psychotherapy to pastoral counseling.

For some time, the mental health professions have been suffering a critical shortage in manpower. The professions of psychiatry and clinical psychology are producing practitioners at an ever-increasing rate, yet the demand continues to outstrip the supply. The mental health manpower needs of the country have been well documented by Albee (1959) and Schofield (1964).

This high demand for therapeutic services and the short supply of practitioners has resulted in the use of as yet untapped resources to help fill the gap between supply and demand. This in turn has led to the establishment of subprofessional groups to carry out and implement some of the mental health practices which the short supply of professionals presently does not permit. Such programs have crystallized around the development of the mental health technician, for example, the work of Holzberg (1963, 1965) and others in the training and use of college students in the mental hospital and that of Rioch, Elkes, Flint, Usdansky, Newman, and Silber (1963) in the training of mature but inexperienced housewives to perform as therapists. Truax and Carkhuff (1967, Ch. 7) have recently reported on a program for the training of lay hospital personnel as therapists.

This approach raises some serious questions about the type of people who seek these opportunities and, in addition, some concern about the levels of empathy that these subprofessionals and novices can provide. In general, one may ask, "What personality characteristics are necessary for a successful helping relationship?" This question has, in part, been investigated in the research of Bergin and Solomon (1963) on some of the test correlates of empathy. They attempted to compare the level of empathy of 18 postinternship students in clinical and counseling psychology (using the Truax Accurate Empathy Scale) with Minnesota Multiphasic Personality Inventory

¹ This report is based on a paper presented at the meeting of the Midwestern Psychological Association, Chicago, May 1967, and in part on a doctoral dissertation submitted to the University of Cincinnati, 1967.

² Also at the Hamilton County Diagnostic Clinics. Requests for reprints should be sent to George J. Vesprani, Hamilton County Diagnostic Clinics, 295 Eikenbrecher Avenue, Cincinnati, Ohio 45229.

(MMPI) and Edwards Personal Preference Schedule (EPPS) profiles. The results they obtained, though based on a relatively small sample, are provocative and partially support other findings with regard to the positive relationship of empathy and adjustment. On the MMPI, for example, they found the Depression (*D*) and Psychasthenia (*Pt*) scales to be negatively and significantly correlated with accurate empathy (AE). Other negative, though not significant, correlations were reported for those scales indicative of personality disturbance. The EPPS data indicated that Consistency (Con), Order (Ord), and Intracception (Int) scales are all negatively and significantly related to empathy, while Dominance (Dom) and Change (Chg) scales correlated positively and significantly with AE.

Perhaps the most significant finding of this study was the reported negative relationship between measures of subjective discomfort (*Pt* and *D* scales of the MMPI) and AE. This suggests that relevant standards for the selection of clinical psychologists or therapists in general should include some measure of these significant personal attributes.

This paper reports an attempt to cross-validate and extend the previous work of Bergin and Solomon (1963) using a group of untrained college students in a companion-like program instead of graduate students in clinical and counseling psychology. Such a program, based on the work of Holzberg (1963, 1965), uses untrained college students who are to act in a quasi-therapeutic relationship with mental patients. In addition, this study attempted both to evaluate the initial level of AE and to explore longitudinally the development of AE over a course of contacts between patients and college companions.

On the basis of the existing research data of Bergin and Solomon, it was hypothesized that companions with high scores on the *D* and *Pt* scales of the MMPI would manifest a low level of AE in their relationship with these patients. A second hypothesis was that those companions with high scores on the Dom and Chg scales of the EPPS would manifest high levels of AE. The final hypothesis was that a longitudinal evaluation would show an

increase in the level of AE as a result of lessened anxiety on the part of the companions and their greater familiarity with the task.

METHOD

To accomplish these goals, students in abnormal psychology classes at the University of Cincinnati were administered the EPPS and the *D* and *Pt* scales of the MMPI. Thirty-three female volunteers were then solicited from these classes to participate as companions to female patients in a local mental hospital. These companions ranged in age from 20-25 years with a median age of 21 years. The previous work of Bergin and Solomon (1963) suggested a negligible relationship between grade average, intelligence (as reflected by the verbal score on the Graduate Record Examination), and AE. For this reason, these data were not gathered or used as selection criteria. The decision to use only female Ss was spurred by the fact that previous studies in this area have used mixed populations and have not controlled for this variable. As companions, these students acted in a quasi-therapeutic relationship with a patient arbitrarily assigned to them for a series of five 1-hour sessions over a 3-week span of time. Patients were selected solely on the basis of their ability to communicate at least at a minimal level as judged by the ward psychiatrist. These patients ranged in age from 17-68 years, with the median age of 36 years. Fourteen of these patients had a schizophrenic diagnosis, usually undifferentiated type, five others were considered psychotic depressives, three had a diagnosis of alcohol or drug intoxication, and the remainder had a variety of neurotic conditions usually with anxiety or depression as a main focus. The hospital in which they were patients is a receiving hospital and usually keeps its patients for 90 days or less.

Training of Companions

Most of the students participating as companions had no prior experience as interviewers. It was felt, therefore, that they should have some acquaintance with the tape recorder, the materials, and the setting prior to their actual interview experience. To accomplish this, the companions were brought together at the hospital for a group meeting. A written and verbal description of their task was presented to them, including the general instructional set to "Get to know the patient as well as possible." They were told to be good listeners as well as interviewers. Their role was to be "companionable, to ask questions which invite the patient's participation, and to comment on what the patient says." In addition to the verbal and written description of their task, they were given a vicarious learning experience. This consisted of listening to several 10-minute taped samples of interactions between students and patients, followed by an opportunity to role play both the patient and companion roles. After this brief training, they were assigned to their patients.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF
MMPI AND EPPS VARIABLES

Variable	Present study (N = 33)		Bergin and Solomon (N = 13) ^a	
	M	SD	M	SD
EPPS				
Con	50.91	9.65	50.36	7.14
Aff	46.48	11.62	46.29	11.02
Def	42.21	9.74	43.43	10.81
Ach	54.54	11.65	56.93	10.78
Ord	46.09	9.26	45.79	10.77
Exh	51.54	9.21	50.93	9.25
Aut	50.42	10.68	51.64	8.67
Int	49.27	12.90	57.29	7.60
Suc	55.09	9.35	54.71	13.10
Dom	48.00	9.02	44.00	12.26
Aba	50.55	11.18	42.00	5.92
Nur	51.24	9.63	49.36	6.51
Chg	47.67	10.91	46.36	10.50
End	46.00	8.47	49.36	10.90
Het	57.21	8.55	56.50	9.63
Agg	52.12	8.94	53.64	6.95
MMPI				
D	21.33	5.16	19.08	4.42
PI	15.03	6.09	9.58	4.41

Note.—T scores for EPPS data and raw scores for MMPI data.

^a N = 12 for MMPI.

Measurement of Accurate Empathy

These sessions with patients were tape recorded and each recording was arbitrarily divided into four segments of equal length. Two, 2-minute verbal samples were then randomly drawn, one from the second segment (early) and one from the fourth segment (late) of each session. Two samples were drawn to ensure a reliable assessment of the level of AE present during the session and, in addition, provide the opportunity to check on change in the level of AE within each session should this occur. To obtain a longitudinal picture, the first, third, and fifth sessions were sampled. Samples were transcribed and presented to three trained judges for rating on the Trust Accurate Empathy Scale.

Empathy in the sessions was measured by applying the 9-point Accurate Empathy Scale previously used by Trust and others to the transcribed samples. Although the typewritten samples contained the conversation of both patient and companion, the scale was designed to characterize distinguishable degrees of AE as represented in the companion's attitudes and responses to the patient. Generally, AE is defined by Trust (1961) as "the sensitivity to current feelings and the verbal facility to communicate this understanding in a language attuned to the client's current feelings (p. 21)." Involved in the procedure for obtaining ratings was the training of three judges

to work with this scale. They, too, were females of similar age and educational background to the companions and had a limited psychological background. They were trained through the use of trial samples before proceeding to the actual data. Once these AE ratings were obtained, they were correlated with the above-mentioned test variables and compared to the Bergin and Solomon data.

RESULTS

The three raters were able to achieve a reliability of .60 for the total data as calculated by use of the intraclass correlation formula described by Ebel (1951). This is a statistically significant level of reliability, identical to that previously reported by Kiesler, Mathieu, and Klein (1966), who used a similar population of raters and essentially identical training procedures.

Table 1 presents the means and standard deviations for MMPI and EPPS variables. The means and standard deviations of EPPS variables have been converted to T scores and, as such, the data from the present study and the Bergin and Solomon study can be directly

TABLE 2
CORRELATIONS OF SUMMED AE SCORES WITH EPPS
AND MMPI VARIABLES IN THE PRESENT STUDY
AND CORRELATIONS FOR THE BERGIN AND
SOLOMON STUDY

Variable	Present study (N = 33)	Bergin and Solomon (N = 13) ^a
EPPS		
Con	-.131	-.54*
Aff	-.081	-.18
Def	.027	-.23
Ach	.046	-.15
Ord	.079	-.41*
Exh	.070	-.12
Aut	.149	.37
Int	.162	-.53*
Suc	-.350**	.00
Dom	.198	.54*
Aba	-.013	-.12
Nur	.118	-.19
Chg	-.340**	.55*
End	.224	.08
Het	-.252	-.05
Agg	-.085	.04
MMPI		
D	-.336*	-.41
PI	-.288*	-.54*

^a N = 12 for MMPI.

* p < .05, one-tailed test.

** p < .05, two-tailed test.

TABLE 3
MEANS AND STANDARD DEVIATIONS OF AE RATINGS

Time	Sessions								Bergin and Solomon	
	1		3		5		Total (1 + 3 + 5)			
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Early	2.90	.80	2.90	.86	2.91	.80	2.90	.83		
Late	2.93	.84	2.78	.81	2.69	.82	2.80	.83		
Total	2.91	.84	2.84	.84	2.80	.82	2.85	.83	2.51	1.01

compared. MMPI means and standard deviations are also reported in Table 1, but as raw scores without *K* correction.^a The *K* correction was not available because the entire MMPI was not administered.

Table 2 presents correlational data for AE scores and test variables. Only the first hypothesis with regard to MMPI variables was verified, as the *D* and *Pt* scales of the MMPI were shown to have significant negative correlations with AE. A correlation of $-.336$ was obtained for *D* and $-.288$ for *Pt*, both of which are significant at the .05 level. The second hypothesis, which predicted that Dom and Chg scales of the EPPS would be significantly and positively correlated with AE, was not confirmed. Although Dom was positively related, it failed to reach significance, and Chg, contrary to the prediction, was significantly but negatively correlated with AE ($-.34, p < .05$).

Inspection of Table 3 indicates that AE did not increase as a function of the number of sessions as predicted in the third hypothesis. Instead, AE scores obtained from samples taken early (second segment) in the session remained remarkably stable over the series of five sessions. This is not true, however, of the AE scores obtained late (fourth segment) in the sessions. The late samples generally tended to yield a noticeably lower level of AE than early samples within each session, and, in addition, this tendency increased as a function of sessions. In effect, the decrease in level of AE between early and late samples is greater in Session III than in Session I and greater

in Session V than Session III. This change is graphically represented in Figure 1.

An analysis of variance indicated that this early-late differential in AE scores failed to reach statistical significance. The differences in total AE scores across sessions as well as the Sessions \times Sample Location interaction were likewise not significant. An attempt was made to further investigate this tendency for late samples to decrease as a function of sessions by means of a trend analysis as described by Edwards (1964). The goal here was to see if this decline over sessions actually defined a negative slope as significantly different from

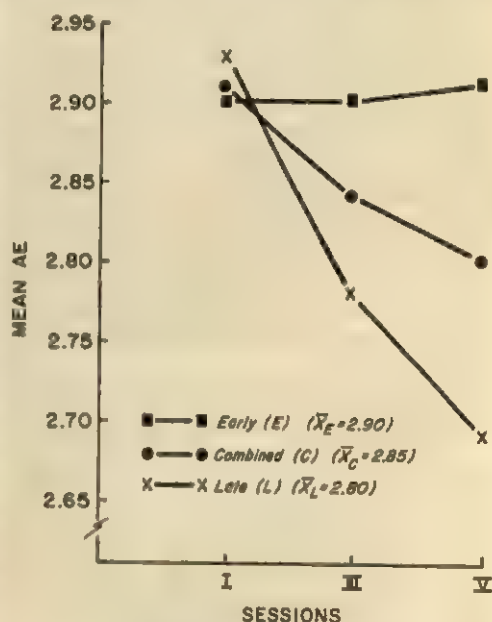


FIG. 1. Mean AE scores for early, late, and combined samples by sessions.

^a Bergin kindly made the raw score data for MMPI scales available so that comparisons could be made.

a horizontal line. This analysis revealed a trend which approached but failed to reach significance at the .05 level. The obtained F ratio was 3.86 and an F ratio of 3.90 is needed for significance at the .05 level.

DISCUSSION

The data with regard to the MMPI variables suggest that the admission of neurotic symptomatology, as indicated by high scores on D and Pt , counterindicates high levels of AE. This was previously reported by Bergin and Solomon (1963) and has now been cross-validated using a much different population in which training level was markedly different and participants were restricted to females. This finding supports the common sense assumption that poorly adjusted people have trouble being empathic although the present data do not suggest why this is the case.

This study has implications for companion-like programs in that it indicates naive female Ss can provide a level of AE at least as high as that of advanced graduate psychology students as reported by Bergin and Solomon (see Table 3).⁴ Companions in this study attained a degree of AE approaching Level 3 on the Truax Accurate Empathy Scale.⁵ The question of whether or not this is an adequate level for notable improvement of the patient relates to the relative usefulness of companion programs. If one considers the level of AE offered by graduate psychology students as adequate to initial progress in therapy, then one must also consider a similar level in companions as acceptable. If this is not the case, that is, if it is felt that a level of AE approaching Level 3 on the Accurate Empathy Scale is not sufficient to produce any patient gains in the therapeutic or quasi-therapeutic context, then

we need to take a look at the training of graduate students who major in clinical areas. It is indeed surprising that naive Ss can manifest a level of AE of similar magnitude to those trained to be therapeutic practitioners. This is even more surprising in that most graduate programs provide the student with extensive therapeutic experience, although without a specific emphasis on empathic understanding per se. Perhaps it is felt that with professional growth and experience, as well as good psychological adjustment on the part of the practitioner, one's empathic abilities are sharpened and do not require specific training. Certainly the level of AE is higher in professionals and it can be assumed that if this level were not a function of training or selection, it must have grown out of experience. Just as one might expect graduate students to improve in level of AE offered to a patient as a function of experience, if for no other reason than the relatively higher level of AE offered by professionals, one might expect an advancement in AE scores of companions with further experience. The degree of improvement and the need for specific training to aid this process is a suitable subject for future research.

No formal evaluation was undertaken to assess the experiential aspects of the study from the companion's viewpoint. Most commented that it was a valuable and interesting experience and that they would continue in such a role if the opportunity were available. This was similar to the patients' verbalized comments that the experience was helpful. The patients filled out a five-item rating scale characterizing their experience. This resulted in a rating set in which the extreme positive rating was chosen most often and may be a reflection of what they felt E expected as opposed to an accurate representation of their investment. This aspect certainly needs further investigation.

The second major finding of this study relates to the methodological problems regarding the sampling of this variable. It has been thought that AE was a stable therapist variable. Indeed, if one samples from the early-to-middle portions of the therapeutic interaction, one obtains a reasonably stable measure of AE. As demonstrated by the present study,

⁴ A one-to-one comparison is difficult if not impossible in that Bergin and Solomon, for the purposes of their analysis, expanded the AE scale to 10 points by inserting an additional point between Empathy Levels 2 and 3.

⁵ Truax (1961) defines Level 3 as:

Therapist responds accurately to client's more exposed feelings. He also displays concern for the deeper, more hidden feelings, which he seems to sense must be present, though he does not understand their nature. The therapist seems to assume the presence of deep feelings, although he does not sense their meaning to this particular patient [p. 8].

however, sampling in the last quarter results in a consistently lower measured level of AE. Not only does the level of AE decrease within the session itself, as a function of time, but the decrease accelerates with each succeeding session. This trend has not been reported before because other researchers have chosen to sample the early segments of the therapeutic interaction. The only previous report of instability in this therapist variable comes from Kiesler et al. (1966), who reported on longitudinal sampling from the last half of the first 15 sessions of the therapeutic interaction. He reported considerable variability in the measured level of AE in the first seven or eight sessions and a stabilization at about the twelfth session.

The meaning of this trend is open to speculation at present, but its implication is that sampling of this variable from therapy interactions should take cognizance of what appears to be a systematic change in AE. Certainly, this stability, or lack of it, has relevance when one attempts to interpret the results of such studies. This suggests (a) that studies might be done taking more as well as standardized samples from each session to ascertain other fluctuations or trends within the session, that is, sample from the same location for all sessions sampled; and (b) a further exploration of longitudinal trends in AE over the full range of therapeutic contact is needed. Most significant of all is the suggestion that care must be taken in interpreting measurement results with regard to sample location.

In summary, the cross-validation of the Bergin and Solomon study, which this study attempted, produced some rather interesting results. Perhaps the most significant finding of this study is the confirmation of the hypothesis regarding the relationship of the *D* and *Pt* scales of the MMPI to AE. The data suggest that neurotic self-descriptions, as indicated by high scores on the *D* and *Pt* scales of the MMPI, counterindicate high levels of AE. This study did not cross-validate the findings of Bergin and Solomon with regard to EPPS variables, suggesting only a superficial and variable relationship to AE. Finally, the second major finding of this study is demonstrated in the tendency for the level of AE

to decline between the second segment sample and the fourth segment sample within each session. In addition, this decrease in level of AE becomes more marked as a function of increasing sessions. This trend seems to imply that sampling of this variable from therapy interactions should take cognizance of what appears to be a systematic change in the level of AE over time.

REFERENCES

- ALBEE, G. W. *Mental health manpower trends*. New York: Basic Books, 1959.
- BERGIN, A. E. The empirical emphasis in psychotherapy: A symposium. The effects of psychotherapy: Negative results revisited. *Journal of Counseling Psychology*, 1963, 10, 244-250.
- BERGIN, A. E., & SOLOMON, S. Correlates of empathic ability in psychotherapy. Paper presented at the meeting of the American Psychological Association, Philadelphia, September 1963.
- EBEL, R. L. Estimation of the reliability of ratings. *Psychometrika*, 1951, 16, 407-424.
- EDWARDS, A. L. *Experimental design in psychological research*. New York: Holt, Rinehart & Winston, 1964.
- HOLZBERG, J. D. The companion program: Implementing the manpower recommendations of the joint commission on mental illness and health. *American Psychologist*, 1963, 18, 224.
- HOLZBERG, J. D. The social interaction of college students and chronically ill mental patients. *American Journal of Orthopsychiatry*, 1965, 35, 487-492.
- KIESLER, D. J., MATHIEU, P. L., & KLEIN, M. H. Measurement of conditions and process variables. In C. R. Rogers, E. T. Gendlin, D. J. Kiesler, & C. B. Truax (Eds.), *The therapeutic relationship and its impact: A study of psychotherapy with schizophrenics*. Madison: University of Wisconsin Press, 1967.
- ROCHE, M. J., ELKES, C., FLINT, A. A., USDANSKY, B. S., NEWMAN, R. G., & SILBER, E. National Institute of Mental Health pilot study in training mental health counselors. *American Journal of Orthopsychiatry*, 1963, 33, 678-689.
- SCHOFFIELD, W. *Psychotherapy: The purchase of friendship*. Englewood Cliffs, N. J.: Prentice-Hall, 1964.
- TRUAX, C. B. A scale for the measurement of accurate empathy. In *Discussion papers*. Psychiatric Institute, Madison: University of Wisconsin Press, 1961.
- TRUAX, C. B. The empirical emphasis in psychotherapy: A symposium. Effective ingredients in psychotherapy: An approach to unraveling the patient-therapist interaction. *Journal of Counseling Psychology*, 1963, 10, 256-263.
- TRUAX, C. B., & CARKHUFF, R. R. *Toward effective counseling and psychotherapy*. Chicago: Aldine, 1967.

(Received October 14, 1968)

IDENTIFICATION OF CLINICALLY RELEVANT DIMENSIONS OF CHILDREN'S BEHAVIOR¹

JACOB O. SINES,² JEROME D. PAUKER

University of Missouri

LLOYD K. SINES

AND

DAVID R. OWEN

University of Minnesota

University of Missouri

The Missouri Children's Behavior Checklist is a set of descriptions of children's behavior that may be rated by a child's parent. This checklist is based on the literature dealing with observers' reports of children's behavior in a number of different settings and consists of items that have been shown to cluster together into six relatively independent dimensions or scales (Aggression, Inhibition, Activity Level, Sleep Disturbance, Somatization, and Sociability). The scales can be rated with reasonable reliability, and the scale scores, based on ratings by mothers, discriminate at least two groups of clinically different children. The dimensions of the Missouri Children's Behavior Checklist appear to be sufficiently independent, reliable, and discriminating to warrant further use as one approach to the quantitative evaluation of the behavior of children.

This report deals with the construction of a checklist (Missouri Children's Behavior Checklist) for the assessment of several relevant dimensions of children's behavior. The primary reason for developing this checklist was to provide a method for identifying groups of children, each of which would be at the extreme of one of several clinically or theoretically significant dimensions of children's behavior. These checklist-identified groups, quantifiably different in described behavior, were meant to serve as criterion groups in an ongoing program of research designed to construct and validate an objective, nonverbal, personality test for children (Sines, Pauker, & Sines, 1966). A by-product of the construction of the checklist has been the development of the several checklist scales which, apart from their use in the definition of criterion groups, have shown promise in themselves of providing useful research and clinical data.

The literature dealing with empirical studies of the behavior of children shows marked consistency in pointing up a number of dimensions that appear to describe the behavior

of children even when those dimensions were derived from rather different populations of children and somewhat different methodologies were used (Dreger, Reid, Lewis, Overlade, Rich, Taffel, Miller, & Fleming, 1964; Jenkins & Hewitt, 1944; Ross, Lacey, & Parton, 1965; Spivack & Levine, 1963-1964). On the basis of this literature, it seems clear that within broad limits, professionals as well as parents and other nonprofessionals use highly similar general language, concepts, and dimensions of behavior to describe children's behavior (Walker, 1967).

Views concerning the number of dimensions that are necessary and sufficient to reflect behavioral differences between children are characterized by less agreement between investigators; some feel that a small number of dimensions can account for the valid variance in reports of children's behavior (e.g., Peterson, 1965) and some have located and tentatively labeled a relatively large number of factors or clusters of specific behaviors (e.g., Dreger et al., 1964; Spivack & Levine, 1963-1964). There are equally important related questions, of course, such as whether behavioral or personality dimensions, however identified and in whatever number, will allow us to improve our prediction of children's behavior and whether they will allow us to discriminate between different etiologies of various patterns of children's behaviors.

¹This work was supported in part by United States Public Health Service Grant MH HD 13657.

²Requests for reprints should be sent to Jacob O. Sines, who is now at Case Western Reserve University, Department of Psychology, Cleveland, Ohio 44106.

As the first step in a larger study of these latter questions, the authors examined the literature for behavior descriptions and clusters of descriptive statements that appeared to relate to behavior dimensions of particular clinical value. Previous investigators generally made use of parents' reports of their children's problem behavior (Dreger et al., 1964) or statements of referral problems or case records in child guidance clinics or other institutions (Jenkins & Glickman, 1946; Peterson, 1961; Spivack & Levine, 1963-1964). Ross et al. (1965) recognized this bias toward pathology and included a set of "good adjustment" items in a "pro-social behavior" checklist scale.

In each of the numerous studies that have been reported (Dreger et al., 1964; Fish & Shapiro, 1964; Patterson, 1964; Peterson, 1961; Quay & Quay, 1965; Ross et al., 1965; Spivack & Levine, 1963-1964; Walker, 1967), children were rated by parents, teachers, institutional supervisors or clinicians either as showing or as not exhibiting or not possessing each of the behaviors or characteristics selected for study. The rating data were typically subjected to some sort of clustering procedure, usually factor analysis, in order to identify those behavior items that tend to occur together.

In many of the studies previously listed, three consistent and relatively independent sets or clusters of behavioral characteristics have been noted. These three factors or clusters have been variously labeled, but may reasonably be referred to as "Aggression," "Inhibition," and "Activity Level."

In addition to these three frequently noted factors, several additional clusters of behaviors have been described, primarily in the work of Dreger et al. (1964). Those of greatest clinical interest here are "Sleep Disturbance," "Somatization," and "Sociability."

The behavior checklist reported in this paper was derived from several of the checklists and behavioral descriptions which appeared in the literature referred to above.

METHOD

The final form of the Missouri Children's Behavior Checklist, consisting of 70 statements, was distilled statistically from 95 behavior-descriptive

statements which made up an original form of the checklist. The original 95 statements were selected from the existing literature to cover six dimensions of behavior: aggression, inhibition, activity level, sleep disturbance, somatization, and sociability. No item contributes to more than one dimension. The individual items were selected from the published reports if they were listed with factor loadings of .30 or greater, if they were employed by the investigator in his narrative description of the behavior dimension, or if they appeared in the same cluster in more than one study.

Data on the original checklist were collected in 15 child psychiatry, pediatric, and mental health clinics in the United States and Canada.³ Most of the children, 404 boys and 250 girls, ages 5 through 16 years, were seen for psychological evaluation. The mother of each of these children was asked to indicate on the checklist (by circling a "Yes" or a "No") whether or not her child had shown the behavior described in each of the 95 statements during the previous 6 months.

The analyses reported below deal with the checklists of the 404 boys. Similar analyses will be made for girls when considerably more data have been collected.

RESULTS

Construction of the Final Checklist Form

A score was derived for each of the six behavior dimensions by assigning one point to the dimension for each of its items to which the mother responded in the direction of the published factor loading. Point-biserial correlations were calculated between each of the 95 original items and the total score on each of the six dimensions or factors. An item was retained in the checklist and assigned to a dimension if the point-biserial correlation between the item and the total dimension score was at least .30, and if the square of the point-biserial correlation was at least twice as large as the square of the correlation between that item and the total score on any of the remaining five factors. Several items were thus eliminated. The checklist items and the direction of scoring are listed in Table 1.

³ The authors wish to express their sincere appreciation to the following psychologists who furnished data on the Missouri Children's Behavior Checklist in connection with this study: Loretta K. Cass, Charles S. Cleeland, Bruce R. Cushna, James W. D. Hartman, Hyman Hops, Jon Krapfl, Luciano L'Abate, Donald S. Leventhal, David Markert, Eileen T. Nickerson, Elizabeth C. Pennick, and Gertrude J. Williams.

TABLE 1
FINAL CHECKLIST ITEMS AND SCORED RESPONSES

A-1	Says as for instance, "I'll get even," "You won't get away with that," "I'll show him," expresses desire for revenge (Y)	H-4	Over-talkative, chatters, keeps talking or interrupting conversation (Y)
A-2	Fights (Y)	H-5	Sings or hums continually (to the expressed annoyance of others) (Y)
A-3	Says "Others are to blame" for own actions (Y)	H-6	Becomes more active or more talkative in groups, becomes noisier and more excited than usual when he is in a group (Y)
A-4	Selfish (Y)	H-7	Speaks rapidly, words "come tumbling out fast" (Y)
A-5	Unscrupulously takes advantage of others (Y)	H-8	Becomes "jittery," building up tension, becomes wound up (Y)
A-6	Hurts animals (Y)	H-9	Is said to be distractable, turns away quickly from what he is doing when something else moves, when someone speaks, or other sounds are made (Y)
A-7	Makes statements contrary to fact (lying, telling untruths) (Y)	H-10	Falls, cuts, bruises, injures self, has many accidents (Y)
A-8	Screams, bangs objects when denied something, has temper tantrums (Y)	Sl-1	Cries out in sleep (Y)
A-9	Pulls other children's hair, punches, steps on toes, etc., annoys children (Y)	Sl-2	Talks in sleep (Y)
A-10	Steals (Y)	Sl-3	Complains of bad dreams (Y)
A-11	Destroys or defaces property (Y)	Sl-4	Tosses and turns in sleep, rolls, gets up often at night, etc. (poor or restless sleeping) (Y)
A-12	Plays with matches (Y)	Sl-5	Talks about or complains of nightmares about past serious events (divorce, automobile accident, fire, loss of loved one, or other "crisis" events) (Y)
A-13	Hurts other children (pinches, hits, kicks or other destructive acts) (Y)	Sl-6	Walks in sleep (Y)
A-14	Teases other children (Y)	Sl-7	Sleeps well, awakes very few times at night (a "good sleeper") (N)
A-15	Hits smaller children, "picks on" weaker or smaller children (Y)	Sl-8	Has difficulty going to sleep (Y)
A-16	Screams more than others (Y)	Sl-9	Irregular bed time (Y)
A-17	Threatens to kill someone (Y)	Ps-1	Becomes so upset by changes in routine such as changing residences or schools or when expecting visitors that the child may vomit or report bodily aches, headaches, stomach aches, or feelings of nausea (Y)
A-18	Swears or curses (uses "Hell," "God damn" or other four-letter words) (Y)	Ps-2	Worries a great deal, is said to be a worrier, expresses worry or concern about bad grades, health, etc. (Y)
A-19	Does not answer when spoken to, pouts, looks mean or sullen (Y)	Ps-3	Clings to mother (stays close to mother, hangs onto dress or hand) (Y)
I-1	Prefers to be with children younger than himself (Y)	Ps-4	Complains of pains in head (Y)
I-2	Is apathetic or underactive (Y)	Ps-5	Complains of pains in limbs or back (muscle aches and pains) (Y)
I-3	Does not try new situations, "hangs back," is considered by others as fearful or shy (Y)	Ps-6	Vomits when things "do not go his way," when he shows signs of anger (red face, raised voice, etc.), when he says he is worried, or when he feels sad or is emotionally upset (Y)
I-4	Does not perform before group, refuses to speak before class when requested, does not volunteer to speak or act before group or class (Y)	Ps-7	Cries at separation from mother (on going to school, camp, etc.) (Y)
I-5	Withdraws, remains quiet, does not talk back when others shove, hit, accuse, or criticize him (does not "stand up for self") (Y)	Ps-8	Has physical complaints (Y)
I-6	Speaks with weak voice, in a monotone, voice "trails off" at ends of sentences, or speaks in a weak, high-pitched voice (Y)	So-1	Expresses or shows concern over the misfortunes of others (e.g., pats shoulder, asks questions about troubles, says "you feel unhappy, don't you?" (Y)
I-7	Stays largely in room or house (Y)	So-2	Expresses appreciation for others' acts (Y)
I-8	Says "I'm tired," "I want to rest," etc. (others say that he tires easily or rests often) (Y)	So-3	Expresses delight over the happiness of others (e.g., claps hands, says "that's good!") (Y)
I-9	Cries easily (Y)		
I-10	Is sensitive (Y)		
I-11	Has few close friendships (Y)		
I-12	Is exclusive, prefers to be by himself (Y)		
I-13	Does not participate in group activities, stays in background (said to be retiring) (Y)		
I-14	Is shy or timid (Y)		
H-1	Moves constantly, "gets into everything," "swarms all over," is overactive (Y)		
H-2	Jumps from one activity to next, does not finish task (others say he has a short attention span) (Y)		
H-3	Stumbles, falls easily, throws clumsily, is awkward (Y)		

Note.—A = Aggression, I = Inhibition, H = Activity Level, Sl = Sleep Disturbance, Ps = Somatization, and So = Sociability.

TABLE 1—(Continued)

So-4 Expresses delight in beauty (Y)	or conversation with adults other than parents (Y)
So-5 Discusses own problems with others (Y)	
So-6 Says "I'm sorry," "Won't you forgive me?" more than others do (expresses great remorse, apologizes repeatedly, cries after hurting or telling untruths or destroying property) (Y)	So-8 Asks to be held or hugged, climbs into lap, etc. (seeks physical expressions of affection) (Y)
So-7 Talks easily with adults, initiates activities	So-9 Sought out by others, others say they like him; among first selected for teams, etc. (Y)
	So-10 Requests praise or approval (Y)

Each new dimension was rescored, and the point-biserial correlation was calculated between each item and each of the total scores.⁴

The means and standard deviations of

⁴ Tables giving (a) the point-biserial correlations between each checklist item and total scores on each behavior dimension and (b) the percentage of mother-father agreement on each checklist item have been deposited with the National Auxiliary Publications Service. Order Document No. 00660 from the National Auxiliary Publication Service of the American Society for Information Science c/o CCM Information Sciences Inc., 909 Third Avenue, New York, New York 10022. Remit in advance \$3.00 for photocopies or \$1.00 for microfiche and make checks payable to: Research and Microfilm Publications, Inc.

scores on each of these six clusters or behavior dimensions were calculated for the total group and for the boys at each age from 5 to 16 years inclusive. These values are presented in Table 2.

Dimensions of behavior such as these may be reasonably expected to meet at least three requirements if they are to warrant further study: (a) the specific behavior items and the factor scores should be reasonably reliable, (b) the dimensions should be relatively independent, and (c) the several dimensions should be differentially present and measurable in clinically different groups of children.

TABLE 2
MEANS AND STANDARD DEVIATIONS FOR CHECKLIST DIMENSIONS
AT EACH AGE AND FOR TOTAL GROUP OF BOYS

Dimension	Age												Total
	5	6	7	8	9	10	11	12	13	14	15	16	
	N												
	8	26	47	57	69	48	38	41	25	24	14	7	404
Aggression													
\bar{X}	3.12	5.38	5.53	6.05	5.58	6.40	6.60	5.32	6.84	4.79	5.78	2.57	5.74
SD	3.22	3.96	4.09	4.39	3.75	5.62	4.62	4.37	4.86	4.99	3.94	1.13	4.43
Inhibition													
\bar{X}	2.62	3.15	4.26	3.79	3.77	4.02	4.79	4.58	5.28	4.58	5.50	3.00	4.16
SD	2.13	2.22	2.93	2.66	2.47	2.42	3.53	3.01	3.91	2.75	3.30	2.24	2.85
Activity Level													
\bar{X}	2.88	4.42	4.83	4.53	4.65	4.71	4.53	4.49	3.44	3.54	2.43	1.71	4.31
SD	1.56	2.74	2.92	2.64	2.50	2.57	2.77	2.81	2.08	2.30	2.10	1.80	2.64
Sleep Disturbance													
\bar{X}	1.25	1.85	2.47	1.67	1.70	2.19	2.45	2.83	2.16	1.79	1.78	2.43	2.08
SD	1.28	2.13	2.61	1.98	2.12	1.98	2.24	2.05	2.10	1.67	1.80	2.15	2.12
Somatization													
\bar{X}	1.00	1.00	1.49	1.04	1.09	1.98	1.29	1.58	1.88	1.25	1.64	.71	1.37
SD	1.07	1.41	1.68	1.31	1.44	1.88	1.39	1.72	1.51	1.33	1.69	.76	1.55
Sociability													
\bar{X}	6.00	4.88	5.08	5.10	4.84	4.88	3.60	4.78	3.92	3.88	3.50	4.00	4.64
SD	1.93	2.55	2.12	2.64	2.41	2.40	2.21	2.36	2.29	2.21	2.93	1.15	2.41

TABLE 3

ODD-EVEN RELIABILITY OF SCORES ON
CHECKLIST DIMENSIONS

Dimension	Pearson r	Spearman-Brown correction
Aggression	.76	.86
Inhibition	.62	.76
Activity Level	.69	.82
Sleep Disturbance	.62	.77
Somatization	.52	.68
Sociability	.50	.67

The published research has dealt with these issues in several respects, and data relating to each of these questions for the final form of the Missouri Children's Behavior Checklist are reported below.

Reliability

Internal consistency. Scores on odd-even halves of each of the six behavior dimensions were correlated for evidence of one type of reliability.

Interjudge agreement. The authors were able to collect behavior checklists from both parents of 47 children. The agreement between mothers' and fathers' ratings on individual items ranged from 53.2% to 93.6%. The average agreement for all the items included in each dimension or scale of the checklist ranged from 68.9% to 93.2% (see Footnote 4).

Interdimension Correlations

The raw total scores on each of the dimensions were intercorrelated using the data

TABLE 4

INTERCORRELATIONS BETWEEN TOTAL SCORES
ON CHECKLIST DIMENSIONS

Dimension	1	2	3	4	5	6
1. Aggression		.19**	.43**	.27**	.21**	-.10*
2. Inhibition			.07	.21**	.29**	-.11*
3. Activity Level				.28**	.21**	.23**
4. Sleep Disturbance					.40**	.01
5. Somatization						.11
6. Sociability						

* $p < .05$.
** $p < .01$.

from the 404 boys. The intercorrelation coefficients are presented in Table 4.

Between-Group Differences

In order to determine whether scores on these several behavior dimensions were congruent with generally accepted clinical experience, the means and standard deviations of checklist scale scores for 24 boys seen in a university child psychiatry clinic were compared with checklist scores for a group of 24 nonreferred boys who had been tested extensively for another study of "normal" children.

The ages of the boys in each of these groups ranged from 6 through 14 years with a mean of 9.88 years and a standard deviation of 2.21 years. The IQ scores ranged from 80 to 134 in the nonclinic group, with a mean of 104.04 and a standard deviation of 11.87. The IQ scores in the clinic group ranged from 80 to 137, with a mean of 103.67 and a standard deviation of 11.95.

The distributions of scores on each of the checklist factors are presented in Table 5,

TABLE 5

COMPARISONS OF CHECKLIST SCORES FOR NORMAL (N)
AND CLINIC (C) BOYS MATCHED FOR AGE AND IQ

Dimension	Group	\bar{X}	SD	t	p
Aggression	N	3.17	2.55	4.40	.002
	C	8.67	4.85		
Inhibition	N	2.04	2.80	2.11	.05
	C	3.96	3.14		
Activity Level	N	2.46	2.02	3.74	.002
	C	5.38	2.72		
Sleep Disturbance	N	.79	1.77	1.64	ns
	C	1.88	2.21		
Somatization	N	.75	.99	1.76	ns
	C	1.29	1.37		
Sociability	N	5.88	2.11	4.09	.002
	C	3.54	2.00		

along with the mean factor or scale scores. There are significant differences between these two groups of boys on checklist scales of Aggression, Inhibition, Activity Level, and Sociability.

DISCUSSION

The point-biserial correlations between each item and each total scale score indicated that each of the checklist items relates relatively exclusively to only one of the six dimensions of behavior. The dimensions themselves are independent for all practical purposes; although a number of the scale intercorrelations (Table 4) are significant beyond the .01 level, the largest correlation (.43 between Aggression and Activity Level) indicates that the two dimensions in question share only about 18% of their variances.

Reliability is at a reasonable level for a behavior rating scale. The odd-even item correlations (Table 3) indicate an adequate degree of internal consistency. In view of the probable fact that fathers and mothers generally see different samples of their children's behavior, a high degree of parent agreement on the checklist is not crucial for reliability considerations. Nonetheless, the data showed moderately high agreement between parents and were consistent with the degree of agreement reported by other investigators (Dreger et al., 1964; Peterson, 1961; Ross et al., 1965; Quay & Quay, 1965). Although no data on interscorer reliability is reported here, it is the authors' experience that high scoring reliability can be regularly achieved.

The distributions of checklist scores in the clinic and nonreferred groups of boys and the significant differences between the two groups in mean scores on Aggression, Inhibition, Activity Level, and Sociability (Table 5) support the notion that several of the clinically obvious differences between these two groups are reflected in their checklist results. It is possible, of course, that such differences in checklist scores may result from factors such as parental attitudes rather than (or in addition to) intrinsic behavioral differences between the children. Had no differences appeared, however, we would have much less confidence in the concurrent validity of the

Missouri Children's Behavior Checklist so far as distinguishing among children is concerned. (Research relating parent variables to checklist scores is in progress.)

The lack of statistically significant differences between the two groups of boys on the Sleep Disturbance and the Somatization scales may be due to one or both of two major factors. First, the very skewed score distributions on both of these scales reduce the statistical power of parametric analysis. Second, it is quite likely that children whose primary complaints involve disturbances of sleep pattern or numerous somatic symptoms are not referred to a child psychiatry clinic, and this selective referral eliminates such children from our psychiatric sample.

Since the distributions of scores on Sleep Disturbance and Somatization for the two groups of boys do not justify the use of parametric statistics, the Kolmogorov-Smirnov two-sample test was also applied and differences again were found to be insignificant. Thus, the dimensions of Sleep Disturbance and Somatization either are not capable of making useful distinctions among children or the selective referral issue may be highly relevant. Significant differences on these two scales between unselected boys and boys seen in a pediatric clinic would support the selective referral hypothesis.

As has been noted (Dawes, 1962; Sines, 1964), the validity and certainly the utility of an assessment device cannot be established or evaluated by demonstrating differences on the diagnostic instrument between clinically different groups of Ss. Even so, the existence of such differences on the checklist is gratifying. If, however, it is to be useful as anything more than a descriptive measure, research must be done using the strategy recently reported by Kulik, Stein, and Sarbin (1968), in which the behavioral and demographic correlates of patterns of checklist scores were determined. Whether patterns of scores on the several dimensions reported here will be valuable as predictors of extrachecklist attributes remains to be explored.

The information now available does suggest, however, that the use of this checklist by mothers does provide technically accept-

able, specific information concerning several clinically relevant dimensions of their children's behavior.

REFERENCES

- DAWES, R. M. A note on base rates and psychometric efficiency. *Journal of Consulting Psychology*, 1962, 26, 422-424.
- DREGER, R. M., REID, M. P., LEWIS, P. M., OVERLADE, D. C., RICH, T. A., TAFFEL, C., MILLER, K. S., & FLEMING, E. L. Behavioral classification project. *Journal of Consulting Psychology*, 1964, 28, 1-13.
- FISK, B., & SHAPIRO, T. A descriptive typology of children's psychiatric disorders: II: A behavioral classification. In R. L. Jenkins & J. O. Cole (Eds.), *Diagnostic classification in child psychiatry*. (Psychiatric Research Rep. 18) Washington, D. C.: American Psychiatric Association, 1964.
- JENKINS, R. L., & GLICKMAN, S. Common syndromes in child psychiatry. I. Deviant behavior traits. *American Journal of Orthopsychiatry*, 1946, 16, 244-261.
- JENKINS, R. L., & HEWITT, L. Types of personality structure encountered in child guidance clinics. *American Journal of Orthopsychiatry*, 1944, 14, 84-94.
- KULIK, J. A., STEIN, K. B., & SARBIN, T. R. Dimensions and patterns of adolescent antisocial behavior. *Journal of Consulting and Clinical Psychology*, 1968, 32, 375-382.
- PATTERSON, G. R. An empirical approach to the classification of disturbed children. *Journal of Clinical Psychology*, 1964, 20, 326-337.
- PETERSON, D. R. Behavior problems of middle childhood. *Journal of Consulting Psychology*, 1961, 25, 205-209.
- PETERSON, D. R. Scope and generality of verbally defined personality factors. *Psychological Review*, 1965, 72, 48-59.
- QUAY, H. C., & QUAY, L. C. Behavior problems in early adolescence. *Child Development*, 1965, 36, 215-220.
- ROSS, A. O., LACEY, H. M., & PARTON, D. A. The development of a behavior checklist for boys. *Child Development*, 1965, 36, 1013-1027.
- SINES, J. O. Actuarial methods as appropriate strategy for the validation of diagnostic tests. *Psychological Review*, 1964, 71, 517-523.
- SINES, J. O., PAUKER, J. D., & SINES, L. K. The development of an objective, nonverbal, personality test for children. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, May 1966.
- SPIVACK, G., & LEVINE, M. The Devereux Child Behavior Rating Scales: A study of symptom behavior in latency age atypical children. *American Journal of Mental Deficiency*, 1963-1964, 68, 700-717.
- WALKER, R. N. Some temperament traits in children as viewed by their peers, their teachers, and themselves. *Monographs of the Society for Research in Child Development*, 1967, 32(6, Whole No. 114).

(Received October 14, 1968)

RELATIONSHIP OF WORK THERAPY TO PSYCHIATRIC LENGTH OF STAY AND READMISSION¹

MARGARET S. BARBEE²

Fort Logan Mental Health Center, Denver

K. L. BERRY

University of Colorado

AND LEO A. MICEK

Fort Logan Mental Health Center, Denver

It was hypothesized that participants in an in-hospital work-therapy program require a shorter period of both intensive and total psychiatric treatment and fewer readmissions to all psychiatric facilities than do nonparticipants. The Ss were Fort Logan Mental Health Center patients randomly assigned to the participant and nonparticipant groups at admission. Contrary to the hypotheses, the results indicated that participants had significantly longer stays in both intensive treatment and total hospitalization than did nonparticipants. Although there was no difference between the groups with respect to readmission to all psychiatric facilities, participants had significantly more readmissions to Fort Logan. It was suggested that work-therapy programs be established in the community in order to avoid contributing to the hospital dependency of participants.

Proponents of work therapy have suggested that meaningful paid work, deliberately provided for the therapeutic benefit of the psychiatric patient and integrated with the other therapeutic activities, may speed the recovery of acute cases and return long-term patients to usefulness. By communicating more realistic expectations of the requirements and satisfactions of a job, work therapy lessens the fear of leaving the security of the institution and strengthens the patients in the community.

The growing interest in and prevalence of work-therapy programs has been well documented with enthusiastic speculation about the potential for rehabilitation of psychiatric patients with various symptomatologies and levels of dysfunction (Denber, 1960; Jones, 1960). Nevertheless, while some efforts have been directed toward evaluations of the changes in attitudes, behaviors, and outcomes

which obtain in patients as a result of participation in work-therapy programs (Hartlage & Hartlage, 1967; Pile & Swanson, 1968; Thompson, 1961), there is little indication of basic empirical research into these programs' efficiency in expediting the return of participants to their communities and reducing their rehospitalization.

This study was undertaken in order to demonstrate the relationship of work-therapy participation to treatment-outcome variables of Fort Logan Mental Health Center patients. Specifically, it was hypothesized that work-therapy participation would reduce (a) the length of intensive institutional treatment, (b) the length of total hospital care, (c) the rate of readmission to any psychiatric institution, and (d) the rate of readmission to Fort Logan.

The Fort Logan work-therapy program operates independently in a small industrial workshop on the hospital grounds. The program staff members not only direct patient groups in their work, but they also contract with local industries for the light assembling and packaging work to be done in the shop. All materials and supplies are provided by the employer firms, who also deliver the raw materials and pick up the finished products. Groups of patients work for 2-hour periods, generally two or three times per week,

¹This study was done in conjunction with a larger project supported in part by Grant MH 0137-04 from the National Institute of Mental Health. Some of the data were also supplied by the Fort Logan Record System, supported in part by National Institute of Mental Health Grant 5-R11-MH-00931-06. The authors wish to express their appreciation to Joanne Fults for her assistance in the manuscript preparation.

²Requests for reprints should be sent to Margaret S. Barbee, Fort Logan Mental Health Center, 3520 West Oxford Center, Denver, Colorado 80236.

and are paid wages from the fees charged to the employers. The rate of pay is based on the group piecework system, that is, each individual shares equally in the earnings of his work, regardless of his contribution to its success.

METHOD

The research sample was randomly drawn from patients admitted to Fort Logan's Adult Psychiatric Division and randomly divided into two groups, those who were assigned to the work-therapy program (participants) and those who were not selected to attend the workshop (nonparticipants). The sampling procedure stipulated that all Ss must have remained in the hospital for a minimum of 20 days and that the participants must have attended the workshop for at least 10 days, while the nonparticipants must not have attended the workshop at all. The Ss who failed to meet these criteria were deleted from the sample, yielding 110 acceptable participant Ss and 149 nonparticipant Ss at the initial phase of the research. Because this refinement and subsequent sample attrition might have destroyed the randomization of the two groups, statistical tests were performed to establish whether or not there were significant demographic and/or psychiatric differences between participants and nonparticipants. The groups were not found to differ significantly in terms of age, sex, level of education, marital status, diagnosis, or severity of diagnosis, either initially or at any of the three follow-up phases. Follow-up interviews were conducted at 3, 12, and 24 months after the patient's first significant upward transfer. A significant upward transfer is contingent on the decision of the respective patients' therapeutic team, which excludes the workshop staff, and results in a move from a more intensive treatment modality to a less intensive modality, for example, from inpatient status to outpatient status or from day care to a halfway house.

Two years after discharge constituted the time period for determining whether readmission occurred for an S. Sample attrition during this period reduced the participant *N* to 73 and the nonparticipant *N* to 84 for the readmission analyses. Speculation concerning the nonparticipants' greater attrition rate, that is, a loss of 44% from in-hospital to readmission

TABLE 1

CONTINGENCY TABLE FOR LENGTH OF INTENSIVE TREATMENT MEDIAN TEST

Group	Fewer than 103 days	More than 103 days	Total
Participants	42	68	110
Nonparticipants	87	62	149
Total	129	130	259

Note.— $\chi^2 = 9.54$, $df = 1$.

TABLE 2

CONTINGENCY TABLE FOR LENGTH OF TOTAL HOSPITALIZATION MEDIAN TEST

Group	Fewer than 166 days	More than 166 days	Total
Participants	41	69	110
Nonparticipants	89	60	149
Total	130	129	259

Note.— $\chi^2 = 11.89$, $df = 1$.

analyses, as compared to the participants' 34% attrition, is presented at a later point in this paper.

It should be noted that the follow-up interviews were primarily used as the data sources for the above-mentioned control analyses and for the collection of information about readmissions to facilities other than Fort Logan. These interviews, however, were not the data sources used in the testing of the remaining three hypotheses. Data concerning the length of initial intensive care stay, total length of hospital treatment, and readmission to Fort Logan after discharge were obtained from the Fort Logan Record System.

RESULTS

As the distributions of total days of intensive treatments and total length of stay were extremely skewed toward the upper range of the distribution, it was necessary to use the median test for the first two hypotheses. The results of the analysis involving length of intensive treatment, shown in Table 1, yielded a median of 103 days for the combined participant and nonparticipant groups. Thirty-eight percent of the participants and 58% of the nonparticipants had fewer than the median number of days of intensive care, indicating that work therapy participants require more intensive treatment than do nonparticipants. This difference is statistically significant, but contrary to the hypothesis.

For the hypothesis involving total length of stay, the results, shown in Table 2, yielded a similar significant reversal from the prediction, as 37% of the participants and 60% of the nonparticipants were treated in some capacity for less than the median 166 days.

The third hypothesis compared participants and nonparticipants on the basis of number of readmissions to any psychiatric facility after discharge from Fort Logan. Fifty-two percent of the participants and 40% of the

nonparticipants were readmitted. This difference is not significant. Table 3 summarizes the chi-square distribution.

When readmissions specifically to Fort Logan are evaluated, the difference is statistically significant, as 46% and 23% of the participants and nonparticipants, respectively, were readmitted to Fort Logan. Table 4 gives these results. This finding is also contrary to the hypothesis.

DISCUSSION

The results of the above analyses failed to support any of the hypotheses and, in fact, strongly suggest that rather than facilitating the return of the patient to the community, work therapy may have inhibited the transition. Briefly summarizing the statistically significant findings, it was learned that patients who participated in work therapy both remained in intensive treatment and maintained their associations with the hospital through outpatient and low-intensity services for a significantly longer period of time than did patients who were nonparticipants. Participants also returned to Fort Logan in significantly greater numbers than did nonparticipants during the 24 months of the follow-up period, while the differential readmission rate to all psychiatric facilities was not significant. The final two results strongly suggest that work-therapy participation fosters institutional dependency, but the absence of complete data on possible readmissions to other facilities makes this conclusion somewhat tentative. However, if one follows this conclusion, it is possible that the nonparticipants may have obtained postdischarge care from other psychiatric facilities from which

TABLE 4
CONTINGENCY TABLE FOR READMISSIONS
TO FORT LOGAN

Group	Readmitted	Not readmitted	Total
Participants	30	35	65
Nonparticipants	19	50	69
Total	49	85	134

Note.— $\chi^2 = 4.23$, $df = 1$.

data retrieval is less efficient, thus accounting for the greater data attrition in this group.

The interpretation of these results requires a consideration of Fort Logan's goal for the work-therapy program. This program was instituted specifically to eliminate the trend toward continued institutionalization which is frequently seen in psychiatric patients who are hospitalized. However, closer examination of the Fort Logan program reveals some elements which may have created rather than lessened attachment to hospital life. First, for many patients, a workshop job not only offered the most regular and well-paid employment which they had ever obtained, but it also met their vocational needs very adequately. Second, the shop supervisors were supportive rather than demanding, the hours were short, and most behavior problems were well tolerated, in marked contrast to a competitive employment situation. Third, free socializing among patients was encouraged, and these new friends may have tended to replace the influence of family and friends in the community. In summary, the provision of inhospital vocational satisfactions, support from supervisors for low-level functioning, and social contacts exclusively among other patients would not seem to be a positive influence in the return of the patient to his community, where memories of past social and vocational failures are many.

TABLE 3
CONTINGENCY TABLE FOR TOTAL READMISSIONS

Group	Readmitted	Not readmitted	Total
Participants	38	35	73
Nonparticipants	34	50	84
Total	72	85	157

Note.— $\chi^2 = 1.66$, $df = 1$.

REFERENCES

- DENBER, H. C. B. Work therapy for psychiatric patients. *Comprehensive Psychiatry*, 1960, 1, 49-54.
- HARTLAGE, L. C., & HARTLAGE, P. L. Practices contributing to successful rehabilitation. *Mental Hygiene*, 1967, 51, 231-234.
- JONES, M. Social rehabilitation with emphasis on work therapy as a form of group therapy.

- British Journal of Medical Psychology*, 1960, 33, 67-71.
- PILE, E., & SWANSON, B. Industrial therapy and changes in self esteem. *American Journal of Occupational Therapy*, 1968, 22, 282-285.
- THOMPSON, J. R. The effects of industrial therapy upon personality and behavior. *Dissertation Abstracts*, 1961, 21, 3170. (Abstract)

(Received October 14, 1968)

Manuscripts Accepted for Publication in the
Journal of Consulting and Clinical Psychology

- Wechsler Object Assembly Subtest and Bodily Concern: A Review and Replication: Sidney J. Blatt,* Bruce L. Baker, and Jay Weiss: Department of Psychiatry, Yale University, 333 Cedar Street, New Haven, Connecticut.
- A Factor Analytic Study of Performance on the Bender-Gestalt: Jack R. Haynes*: Department of Psychology, North Texas State University, Denton, Texas 76203.
- Achievement Performance Related to Achievement Motivation and Test Anxiety: Joel M. Kestenbaum and Bernard Weiner*: Department of Psychology, University of California, Los Angeles, California 90024.
- Skill versus Chance Activity Preferences as Alternative Measures of Locus of Control: An Attempted Cross-Validation: Juris I. Berzins,* Wesley F. Ross, and Daniel I. Cohen: Department of Psychology, University of Kentucky, Lexington, Kentucky 40506.
- Behavior Ratings of Chronic Hospitalized Patients: Cross-Situational and Cross-Rater Agreement: Julian Rappaport* and Jack M. Chinsky: Department of Psychology, University of Illinois, Champaign, Illinois 61820.
- Further Validation of the Personal Orientation Inventory: Assessment of Self-Actualization of School Counselors: Edwin W. McClain*: Department of Educational Psychology and Guidance, University of Tennessee, Knoxville, Tennessee 37916.
- A-B Distinction and Therapeutic Interaction: Bernard Segal*: Department of Psychology, University of Rhode Island, Kingston, Rhode Island 02901.
- Effects of Different Explanations of Disordered Behavior on Treatment Referrals: Charles E. Colson*: Psychology Department, Institute for Juvenile Research, 907 South Wolcott Avenue, Chicago, Illinois 60612.
- Sample and Personality Biases of Volunteer Subjects: Donald W. Tiffany,* James R. Cowan, and Elaine Blinn: Institute for Community Studies, 301 East Armour Boulevard, Kansas City, Missouri 64111.
- Influence of Client Pretesting on Counseling Outcome: Richard F. Haase* and Allen E. Ivey: Counseling Center, University of Massachusetts, Amherst, Massachusetts 01003.
- Sexual Responsivity of Defensive and Nondefensive Sensitizers and Repressors: Thomas Schill,* Gary Emanuel, Virginia Pedersen, Larry Schneider, and Dale Wachowiak: Department of Psychology, Southern Illinois University, Carbondale, Illinois 62901.
- Identification of Potentially Creative Persons from the Adjective Check List: George Domino*: Department of Psychology, Fordham University, Bronx, New York 10458.
- Personality Correlates of Undergraduate Marijuana Use: Robert Hogan,* Donald Mankin, John Conway, and Sherman Fox: Department of Psychology, Johns Hopkins University, Baltimore, Maryland 21218.
- Empathy and the Reported Emotional Experiences of Beginning Psychotherapists: Jefferson M. Fish*: Psychological Services, State University of New York at Stony Brook, Stony Brook, Long Island, New York.
- Orientations to Seeking Professional Help: Development and Research Utility of an Attitude Scale: Edward H. Fischer* and John LeB. Turner: Department of Psychology, Connecticut Valley Hospital, Middletown, Connecticut 06457.
- A/B Typing and Patient Socioeconomic and Personality Characteristics in a Quasi-Therapeutic Situation: Anne Smith Anzel*: Department of Psychology, University of California, Los Angeles, California 90024.
- Immediate Effects of Self-Confrontation on the Self-Concept: G. Nicholas Braucht*: Colorado State University, 50 West 5th Avenue, Denver, Colorado 80204.
- Alteration of Some Body Image Indexes in Schizophrenics: Joel A. Darby*: Bay Area Child Guidance Clinic, 609 South Monroe Street, Bay City, Michigan.
- Further Investigation of the Mini-Mult: Patricia Brilliant Lacks*: Malcolm Bliss Mental Health Center, 1420 Grattan, St. Louis, Missouri 63104.
- Contextual Constraint and Schizophrenic Language: Ida P. Truscott*: Rollman Psychiatric Institute, 3009 Burnet Avenue, Cincinnati, Ohio 45219.

* Asterisk indicates for whom address is supplied.

SOME PERSONALITY AND APTITUDINAL CORRELATES OF THE "A-B" THERAPIST SCALE¹

JAMES E. DUBLIN,² CHARLES F. ELTON, AND JURIS I. BERZINS

University of Kentucky

Prior research with the "A-B" scale has shown intriguing similarities between practicing therapists and undergraduates varying in A-B status. On that basis, 12 scales from the Omnibus Personality Inventory and the four component scores of the American College Testing Program battery were related to A-B status of undergraduates. A stepwise discriminant analysis showed that the present set of measures significantly discriminated male As from ABs (middles) and Bs, and that the differences between As and Bs involved largely the joint effects of masculinity-femininity, verbal aptitude, and natural science aptitude measures. As and Bs also appeared to differ in a way consistent with previous research on "cognitive style." Correlational and factor-analytic data indicated, however, that the A-B variable cannot be regarded as linearly related to those measures which discriminated As from Bs. A quasi-typological interpretation of the A-B variable was proposed and methodological implications were drawn.

The manner in which therapists respond to items concerned with manual, technical, or mechanical activities on the A-B scale, a short scale derived from the male form of the Strong Vocational Interest Blank (SVIB), apparently relates to differential psychotherapeutic results with certain nosological groups. A series of studies by Whitehorn and Betz (reviewed by Betz, 1967) has shown that therapists who disavow such interests ("As") have higher improvement rates with schizophrenics than those who endorse these interests ("Bs"). The B-type therapists, conversely, achieve significantly better outcomes than As with neurotic outpatients (McNair, Callahan, & Lorr, 1962).

In recent years, no clinical studies comparable in scope to the early studies of Whitehorn and Betz and McNair and associates have been conducted. However, research with the A-B variable has been brought into the laboratory, and a series of therapy analogue studies (reviewed by Carson, 1967) has been conducted. Interestingly, most of these studies have yielded results consistent with the original studies in that, for as yet poorly understood reasons, untrained male Ss,

when asked to interact "helpfully" with simulated patients (e.g., tape-recorded schizoid and neurotic communications), show performance differences consistent with those of practicing therapists. In a variety of analogue situations, A-type Ss have responded "better" than Bs to schizoid, disaffiliative, extrapunitive "others," while Bs have responded more positively than As to neurotic, trusting, intro-punitive "others" (Berzins & Seidman, 1968; Carson, Harden, & Shows, 1964; Sandler, 1965). Such similarities between therapists and untrained Ss strongly suggest that the A-B scale measures a number of basic dimensions of personality.

The empirical derivation of the scale (initially 23 items of the male form of the SVIB discriminated As from Bs) renders its conceptualization difficult, and to date the search for correlates of the A-B variable has yielded suggestive but limited information. Whitehorn and Betz initially observed differences in "clinical style" between As and Bs, and speculated that the interest patterns suggested by scale items may have a direct bearing on therapists' attitudes toward and behavior with patients (Betz, 1967). McNair et al. (1962) favored an explanation in terms of socioeconomic and/or social class differences between their samples and those of Whitehorn-Betz as these might interact with therapists' interest patterns. More recently, Lorr and McNair (1966) observed that since

¹ The authors are deeply indebted to Harriet Rose, University of Kentucky Counseling and Testing Service, and Thomas Greenland, for their cooperation in the administration of the test battery.

² Requests for reprints should be sent to James E. Dublin, who is now at St. Albans Psychiatric Hospital, Radford, Virginia 24141.

A-B scale items correlated significantly with the masculinity-femininity (MF) scale of the SVIB, sex-role identification variables might underlie the differential performances of As and Bs, the former seemingly being less masculine than the latter. Still more recently, Campbell, Stevens, Uhlenhuth, & Johansson (1968) fitted A and B doctors into a hierarchy of occupations according to their mean score on a new Strong-type A-B scale. They found that the A doctor, author-journalist, lawyer, artist, and librarian, in that order, represented scores at one extreme (the A end of their standard scale, the A doctor being most extreme), while (again from most to least extreme) carpenter, pilot, veterinarian, math-science teacher, B doctor, and engineer were at the other extreme. Asserting that their data "clearly support the original Whitehorn-Betz interpretation," Campbell et al. also presented evidence that with respect to A-B scale scores, medical students do not change appreciably during the 4 years of medical school, and suggested that this pattern of scores reflects "an intellectual-verbal versus practical-mechanical dimension." As Carson (1967) points out, the content of the scale defies conceptualization in terms seemingly relevant to psychotherapy (he reports a factor analysis of the items of the scale which showed only that As are uninterested in manual and mechanical activity while Bs show engineering interests). Thus, in summary, the speculations of the above investigators have not received direct empirical test.

The absence of a large-scale attempt simultaneously to investigate a variety of possible personality and aptitudinal correlates of the A-B variable gave rise to the present study. Since previous research has shown intriguing continuities between practicing therapists' performances with patients and untrained Ss' performances with simulated patients, the use of a large sample of undergraduates seemed justified.

METHOD

Subjects. The Ss were drawn from all incoming freshman students who, as part of the ongoing institutional research program at the University of Kentucky, volunteered to take a battery of tests. Over 1,200 males and almost as many females completed the measures pertinent to this study.

Measures. Designed specifically for research with college students (Center for the Study of Higher Education, 1962), the Omnibus Personality Inventory (OPI) was a logical choice as the basic instrument for measuring a number of personality dimensions. The OPI has typically yielded five rotated factors which account for about 80% of the total variance (e.g., Elton & Rose, 1967). However, in the present study, because the authors' interest was unlike that of researchers into problems of academic achievement, the relationship of Ss' A-B status to discrete OPI scales (rather than OPI factors) was examined.

The local version of the OPI (Form C), which excludes four scales because of item overlap, retains the following 12 scales: Thinking Introversion (TI), Theoretical Orientation (TO), Estheticism (Es), Cognitive Complexity (Co), Autonomy (Au), Impulse Expression (IE), Schizoid Functioning (SF), Social Introversion (SI), Religious Liberalism (RL), Masculinity-Femininity (MF), Lack of Anxiety (LA, derived from the Taylor MA scale), and Yea-Nay-sayer (Y-N, derived from the Couch-Keniston scale).³

For measures of scholastic aptitude, the American College Test (ACT) Battery was employed. Its four component scales are Verbal (V), Numerical (N), Social Science (SS), and Natural Science (NS). The above 16 scales constituted the total set of variables to which the A-B scale scores of Ss were related.

The version of the A-B scale used here comprised 19 items (14 from the SVIB and 5 from the MMPI). This scale is derived from item-analytic work by Schiffman, Carson, and Falkenberg⁴ on the 31-item Kemp (1966) modification of the items used in the original Whitehorn-Betz study. Thirteen of the items used in this study are identical to those published by Lorr and McNair (1966) in a 15-item scale which showed an internal-consistency coefficient of .91. The present scale was scored in the "B" direction (high scores = B status).

Selection of groups. Although prior research with the A-B variable has been limited to male Ss, to learn whether the implications of prior research might be extended to females in future studies, extreme and middle groups were drawn from the distribution of A-B scores for both sexes. For males, the mean A-B score was 9.13, $SD = 3.57$. Cutting scores of 5 and 13, one standard deviation to either side of the mean, yielded 98 As and 99 Bs, and 99 Ss whose scores fell between the cutting scores were randomly drawn to comprise the middle (AB) group. For females, the mean A-B score was 5.92, $SD = 3.39$. Similarly derived cutting scores (2 and 10) yielded 98 As and 73 Bs, and 99 ABs were selected as for male Ss.

Analyses. The basic analysis of group differences on the 16 experimental scales consisted of a step-

³ For full details of the meaning of each of the OPI scales and factors, see the OPI Manual. Descriptions of the scales are given only where pertinent to the text.

⁴ Personal communication from Robert C. Carson to the third author, October 1967.

TABLE 1

OPI RAW SCORES AND ACT SCORES FOR MALE AND FEMALE AS, ABS, AND BS

Variable	Male						Female					
	A		AB		B		A		AB		B	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
TI	30.66	9.66	30.20	9.72	28.87	8.96	27.72	7.86	29.08	9.56	30.85	10.29
TO	15.76	4.22	17.26	5.09	17.11	4.18	12.99	4.26	14.65	4.38	16.63	4.77
Es	9.63	5.31	10.34	5.10	10.45	4.41	11.48	4.45	11.74	4.46	13.60	4.52
Co	11.74	4.53	12.71	4.96	13.05	4.20	10.38	3.82	11.29	4.26	12.44	4.62
Au	21.37	8.74	21.27	6.84	19.76	6.44	18.94	6.20	20.17	5.81	19.01	5.82
IE	35.79	9.64	37.46	10.87	37.43	10.32	30.83	9.62	31.41	10.41	32.60	10.41
SF	36.90	11.82	35.44	12.56	35.02	15.33	37.34	13.36	37.09	14.35	34.50	14.40
SI	24.53	8.48	23.55	7.82	24.83	8.21	23.12	10.48	22.93	8.79	22.20	8.13
RL	12.99	5.02	13.09	5.43	13.10	5.65	11.40	4.11	12.20	4.14	11.26	4.65
MF	52.04	7.93	54.98	9.64	55.23	8.76	44.23	9.85	45.45	8.41	47.31	8.70
LA	10.12	4.06	11.11	4.56	11.65	4.38	9.71	4.45	10.67	4.33	10.26	4.31
Y-N	9.46	3.20	9.72	3.63	9.76	3.64	9.81	3.12	9.53	3.33	9.45	3.39
V	21.50	3.41	20.93	3.68	19.98	4.38	22.09	3.62	22.39	3.47	21.38	4.10
N	24.89	5.14	24.53	5.70	24.57	6.22	21.03	5.50	22.83	5.84	22.14	5.47
SS	24.05	5.03	23.30	4.86	23.17	5.45	21.98	5.32	23.20	4.87	23.16	5.42
NS	25.20	4.34	25.01	4.69	25.69	5.37	21.27	5.11	22.32	5.48	22.38	5.27

wise discriminant analysis (Computer Program BMDO7M, version of September 1, 1965, developed by the Health Sciences Computing Facility, University of California at Los Angeles) that describes the rank order in which and the extent to which a number of dependent variables (here, 12 personality and 4 aptitude scales) contribute, individually and jointly, toward a separation of Ss with respect to one or more independent variables (here, the three groups of scorers on the A-B scale).

A second analysis of group differences on the 16 experimental scales (using the same Ss) was a principal axis factor analysis of each criterion group, separately for each sex.

In the factor analysis described above, the number of Ss in each criterion group did not meet the usual rule-of-thumb for this type of factor analysis (10 times the number of variables). Moreover, the groups used in both of the above-described analyses under-represent midrange scorers (ABs). For these reasons, smaller random samples of Ss (117 males and 117 females) were also drawn for correlational purposes. Intercorrelations of the measures were then factored through a principal axis analysis (with A-B score included as the seventeenth variable), rotated to the "normal" criterion proposed by Kaiser (1958).

RESULTS

Mean scores for male and female Ss on each of the 12 OPI and 4 ACT scales are shown in Table 1. As might be expected from consideration of stereotypes in our culture and prior OPI studies regarding sex differences

(cf. Ross, 1968), it is apparent that males have consistently higher scores on MF, IE, TO, N, and NS, while females score higher on Es and V. It appears that for both sexes ABs and Bs have progressively higher scores on masculinity than As, and that for females the same kind of relationship holds for TO. Thus, despite the obviously greater variability of the middle group (ABs), these mean scores would imply a linear relationship between MF and A-B status for males and between TO and A-B status for females.

In the stepwise discriminant analysis, the first discriminant function, which accounted for 83% of the variance, significantly separated the male groups (χ^2 approximation of the Wilks-Lambda test = 49.85, $df = 32$, $p < .025$). Subsequent discriminant functions were not significant by this test. Table 2 shows the order in which and the extent to which each of the 16 variables discriminated the three groups of A-B scale scorers for both sexes. The individually discriminating scales for males (MF, V, NS, and Co) and females (TO and Co) suggest that differentiations on the A-B variable are accomplished largely by scales which show large differences between sexes, and which, as al-

TABLE 2

ORDER IN WHICH AND EXTENT TO WHICH *F* VALUES FOR 16 VARIABLES DISCRIMINATE AS, ABS, AND BS

Order	Males		Females	
	Variable	<i>F</i>	Variable	<i>F</i>
1	MF	3.99*	TO	14.04**
2	V	3.71*	Au	2.45
3	NS	5.11*	Es	2.27
4	Au	1.93	SF	1.84
5	Co	3.10*	Co	2.92*
6	SF	2.43	V	1.63
7	IE	1.72	N	2.09
8	SI	1.47	LA	1.47
9	TO	.93	MF	1.48
10	TI	1.27	SS	1.07
11	LA	.76	RL	.91
12	Es	.52	IE	.52
13	N	.10	Y-N	.55
14	Y-N	.10	TI	.21
15	RL	.10	NS	.20
16	SS	.07	SI	.12

Note.—*df* for males = 2/293; for females = 2/267.**p* < .05.***p* < .001.

ready noted, call to mind sex-role stereotypes and sex differences in interests and aptitudes. The scale which best discriminates the female groups, TO, can in fact be regarded as an MF scale; the scale description ("Interest in science and scientific activity, and a logical, rational, critical approach to problems") clearly suggests masculine pursuits. Moreover, the mean values for the only other scale which individually discriminates the female groups (Co, see Table 1), like the means for MF, appear to be ordered linearly, with both sexes scoring progressively higher on Co as one reads across from As to ABs to Bs. The description of this scale (in brief, "more liking for and tolerance of new, novel, or diverse events, uncertainty and ambiguity") suggests a cognitive style. Therefore, it would appear that the masculinity-femininity variable, in its several manifestations, forms the dominant discriminator of groups drawn from the distribution of A-B scale scorers, and it would appear secondarily that cognitive style differences may be of some importance as they relate to sex-role differences and stereotypes.

Since four measures in this battery individually discriminated the male criterion groups, a further breakdown of these differ-

ences is indicated. Table 3 shows the comparisons of the three male groups on all scales. It is evident that with successive steps in the discriminant analysis, differences between As and Bs carry more weight than A-AB or AB-B differences. The significance of differences between As and Bs reaches its maximum on the third step, that is, after MF, V, and NS have been jointly entered into the discriminant function; subsequent steps show progressively lower *F* values (see Note, Table 3). Recalling that on an absolute basis, male As and Bs differ only slightly with respect to V and NS scores (Table 1), it now appears that in conjunction with MF differences, these differences are quite important in differentiating extreme As from extreme Bs. It is consistent with Bs scoring higher on measures of masculinity that they also score higher on natural science abilities while As outscore them on verbal abilities, and that only slight differences here discriminate the two extreme groups. While Campbell et al. (1968) have speculated about the importance of an intel-

TABLE 3

MULTIVARIATE *F* MATRIX SHOWING COMPARISONS OF A-AB, A-B, AND B-AB DIFFERENCES FOR MALES ON THE 16 VARIABLES

Variable	<i>df</i>	<i>F</i> value		
		A-AB	A-B	B-AB
MF	1/293	5.48	6.47	.04
V	2/292	3.24	6.92	1.49
NS	3/291	3.25	7.58	2.21
Au	4/290	3.17	6.46	1.65
Co	5/289	2.66	6.44	1.90
SF	6/288	2.24	6.14	2.09
IE	7/287	1.93	5.73	2.11
SI	8/286	2.02	5.26	1.84
TO	9/285	1.85	4.71	1.85
TI	10/284	1.67	4.49	1.82
LA	11/283	1.64	4.17	1.65
Es	12/282	1.52	3.91	1.54
N	13/281	1.41	3.59	1.42
Y-N	14/280	1.30	3.34	1.33
RL	15/279	1.22	3.11	1.24
SS	16/278	1.14	2.90	1.17

Note.—Due to the progressively increasing degrees of freedom in the numerator in a stepwise discriminant analysis, several of the A-AB differences are significant at the .05 level, and all A-B differences are significant at the .01 level. However, it is apparent that only the values for the first three scales (MF, V, and NS) increase as each step of the discriminant analysis is carried out, and that this holds true only for the A-B differences.

lectual-verbal versus practical-mechanical difference, here it is demonstrated for the first time on conventional measures of aptitude.

The results of the factor analysis for criterion groups were generally disappointing, but did yield one significant negative finding. The first factor, which uniformly accounted for about 22% of the variance for both sexes and for all criterion groups, was comprised of scale loadings in the high-80s and low-90s on IE, SF, and Y-N.⁵ The failure of the Couch-Keniston measure of acquiescence to discriminate the A-B scale groups is an important negative result in that the items of the A-B scale are keyed largely in the same positive direction, that is, to answer in the "B" direction almost invariably requires saying "yes" or "like." The absence of systematic trends here argues for the relative independence of the A-B scores from the acquiescent response set. (The third author also has unpublished data which reveal that A-B scores are largely independent of the social desirability response tendency as measured by the Marlowe-Crowne scale.)

In view of the findings from the discriminant analysis already discussed, the results of the principal axis factor analysis which included A-B scale score as one of 17 variables were surprising.⁶ In the case of the male Ss, MF loaded (.89) with LA (.43) and on the fourth factor, which accounted for very little (about 9%) of the common variance. A-B did not load with any other scale measure at a level even approaching significance (highest loading was .11, on a factor showing V, NS, and SS loadings ranging from .70 to .84). A sixth factor, which accounted for about 6% of the variance, was comprised of a .90 loading on A-B scale score, but no other scale loaded significantly on this factor. The simple correlation between A-B and MF scale scores, for males, was $-.11$ (ns). Similarly for females, A-B loaded alone (.91) on sixth factor. The highest MF loading (.93) was again on a factor with only two other appreciable loadings (LA, .43 and SI, .51), and the TO scale loaded (.78) on an essentially uninterpretable

third factor (12% of the variance) on which the A-B loading was only .11.

DISCUSSION

Considered in conjunction with the findings of the discriminant analysis, the fact that the A-B variable is minimally correlated with the variables involved in the discrimination of extreme scorers would seem to argue for a quasi-typological interpretation of the A-B variable. That is, while in some absolute sense Bs may be more masculine, less verbal, etc., than As, these differences are not reliably manifested throughout the middle range of the distribution of A-B scale scorers, but only at the extremes (the outer 7% or so).

The conclusions to be drawn from this research should be tempered by noting that the Ss in the present study, unlike practicing therapists and Ss drawn on a voluntary basis to participate in psychotherapy analogue studies, may be more heterogeneous with respect to their intended occupations or professions, and may represent a wider spread of aptitude. For this reason, the results of the discriminant analysis should be understood conservatively. Even so, the MF-V-NS triad, as here measured, can be regarded as important discriminators between extreme As and extreme Bs, in the only study thus far which has attempted both a measure of aptitude and a measure of masculinity-femininity (other than the MF scale of the SVIB itself). That the discriminatory power of these variables applies largely to extreme scorers does not seriously detract from the conclusions which may be drawn from this study. In this connection, it should be noted that in the study of Campbell et al. (1968) already discussed briefly, the A doctor was the most extreme scorer in the A direction while the B doctor was less extreme only than some occupations which are easily conceptualized as primarily mechanical-practical (e.g., carpenter, pilot). There are thus other indications that a quasi-typological rather than continuous relationship of the A-B variable to other measures is tenable.

The direct (experimental) investigation of differences between As and Bs has shown replicable results in only two areas. B-type therapists (Pollack & Kiev, 1963) and under-

⁵ The complete varimax matrices for the criterion group factor analyses are available on request from the first author.

⁶ The complete factor analysis is available on request from the first author.

graduates (Shows & Carson, 1965) appear to exceed As in what Witkin has termed "psychological differentiation" (Witkin, Dyk, Fatterson, Goodenough, & Karp, 1962). That is, Bs were more "field-independent" on the rod-and-frame test than As, who in turn were significantly more variable in performance. Other recent studies suggest that As and Bs may react differentially to stress, with As being inclined to inhibit anger or turn it inward, while the stress reactions of Bs are less well understood (Berzins, Friedman, & Seidman, 1969; Sandler, 1965; Welch & Berzins, 1968).

The findings of the present study can be related, in terms of both theory and methodological implications, to both of these areas of experimentation. Based on the above-described rod-and-frame studies and other correlates of "mode of spatial orientation," Silverman (1967) holds that a "greater communication capacity" of As with schizophrenics "may be conceptualized in terms of perceptual, cognitive and personality variables." He elaborates, however, almost exclusively in perceptual terms:

The A-type . . . is responsive to more stimulus attributes of the perceptual field including incidental social behavior cues. He is more receptive to the effects of seemingly irrelevant stimulation, and to changes in the organization of the perceptual field. He is more capable of relaxing his orientation to reality and responding to hunches and intuition. The A therapist is, thus, more accepting of the "reality" of the schizophrenic patient's perceived unreality. . . . Overall, the perceptual responses of the A-type therapist are more similar to those of the schizophrenic patient than are those of the B-type. The dominant perceptual tendency of the B therapist is to counteract stimulus effects which interfere with articulated, reality-attuned cognitive activity [p. 12].

In view of the above formulation, an important finding is that the outer extremes of the A-B scale score distribution show differences with respect to Co. This is at least consistent with direct experimental evidence suggesting that Bs have a cognitive style which differs from that of As in that it reflects greater "psychological differentiation." To whatever extent a rational, critical, logical approach to problems (TO for females) and the ability to withstand and persevere in the face of new, novel, uncertain, or ambiguous events (Co for both males and females) is similar to "psychological differentiation," there is in the present

study a strong trend for Bs to manifest these characteristics to a greater extent than As. However, if As overall perceptual and cognitive modes are more similar to the schizophrenic patient's than the perceptual and cognitive modes of Bs, as Silverman suggests, then As should have higher scores than Bs on SF, which was not the case. Moreover, if As are "more receptive to changes in the organization of the perceptual field," as Silverman holds, they should be more, not less, "cognitively complex" on a measure such as the Co scale of the OPI. While it is easy to speculate about perceptual, cognitive, and personality differences which are consistent, results such as those of this study discourage broad speculation in the absence of solid empirical referents.

The masculinity-femininity findings of this experiment are generally consistent with the findings of Berzins et al. (1968), Sandler (1965), and Welch and Berzins (1968) that As are inclined to inhibit anger or turn it inward. The B, more masculine, should be more defensive about admitting to feelings of inadequacy or anxiety, and in the factor analysis, MF did load with LA, a scale described precisely in those terms. The dearth of correlates of the A-B scale with variables in the present battery which in principle should relate to differential stress reactions of college-age males (e.g., TI, SI, and especially SF and IE) is consistent with other recent findings that personality scales often do not consistently relate to operationally defined variables except at the extremes of the distribution of scale scores. For example, the failure of middle groups of scorers on the Taylor Manifest Anxiety scale to perform consistently under stress or show "appropriate" physiological reactions is well known, and Dublin (1968) and others have recently shown that the Repression-Sensitization (R-S) scale may not represent measurement along a unidimensional continuum from "repression" to "sensitization."

Summarizing, the absence of correlational findings which demonstrate linear relationships between the A-B variable and the other measures used here and the absence of personality scale findings which would be more consistent with clinical speculations based on

perceptual and cognitive style research require the findings of this study to be understood conservatively. Nevertheless, masculinity-femininity, in its various manifestations, and cognitive style as implied by at least two manifestations (problem solving approach; tolerance for change in the perceptual-cognitive field) do help to discriminate extreme scorers on the A-B scale. Important in this "cognitive style distinction" are verbal-intellectual skills versus practical-nonverbal ones, for even slight differences here differentiate scorers at the extreme of the A-B scale.

Our findings seem important for two differing but related methodological reasons. First, they underscore the importance of middle groups in personality, cognitive, and perceptual research in general, especially when an attempt is made to relate scores on a personality scale to other measures of personality or cognitive or perceptual style. Second, they may show that in personality research, the "phenotype" may bear closer scrutiny than the "genotype." That is, the future of the A-B line of investigation may lie in a continued search for behavioral differences between As and Bs in dyadic interactions involving, or at least modeled after, the kinds of interactions involved in actual psychotherapy, and may not lie in further attempts to refine or extend the A-B scale or to achieve construct validation by building up a nomological network through broad inferences from other closely related personality research areas. Such differences in dyadic behavior may then be understood on other, largely phenotypic, grounds, such as verbal and non-verbal communicative styles.

REFERENCES

- BERZINS, J. I., FRIEDMAN, W. H., & SEIDMAN, E. Relationship of the A-B variable to patient symptomatology and psychotherapy expectancies. *Journal of Abnormal Psychology*, 1969, 74, 119-125.
- BERZINS, J. I., & SEIDMAN, E. Subjective reactions of A and B quasi-therapists to schizoid and neurotic communications: A replication and extension. *Journal of Consulting and Clinical Psychology*, 1968, 32, 342-347.
- BETZ, B. J. Studies of the therapist's role in the treatment of the schizophrenic patient. *American Journal of Psychiatry*, 1967, 123, 963-971.
- CAMPBELL, D. P., STEVENS, J. H., URLEHNUTH, E. H., & JOHANSSON, C. B. An extension of the Whitehorn-Betz A-B scale. *Journal of Nervous and Mental Disease*, 1968, 146, 417-421.
- CARSON, R. C. A and B therapist "types": A possible critical variable in psychotherapy. *Journal of Nervous and Mental Disease*, 1967, 144, 47-54.
- CARSON, R. C., HARDEN, J. C., & SHOWS, W. D. A-B distinction and behavior in quasi-therapeutic situations. *Journal of Consulting Psychology*, 1964, 28, 426-433.
- CENTER FOR THE STUDY OF HIGHER EDUCATION. *Omnibus Personality Inventory research manual*. Berkeley: Author, 1962.
- DUBLIN, J. E. Perception of and reaction to ambiguity by repressors and sensitizers: A construct validity study. *Journal of Consulting and Clinical Psychology*, 1968, 32, 198-205.
- ELTON, C. F., & ROSE, H. A. Traditional sex attitudes and discrepant ability measures in college women. *Journal of Counseling Psychology*, 1967, 14, 538-543.
- KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, 23, 187-200.
- KEMP, D. E. Correlates of the Whitehorn-Betz A-B scale in a quasi-therapeutic situation. *Journal of Consulting Psychology*, 1966, 30, 509-516.
- LORR, M., & MCNAIR, D. M. Methods relating to evaluation of therapeutic outcome. In L. A. Gottschalk & A. H. Auerbach (Eds.), *Methods of research in psychotherapy*. New York: Appleton-Century-Crofts, 1966.
- MCNAIR, D. M., CALLAHAN, D. M., & LORR, M. Therapist "type" and patient response to psychotherapy. *Journal of Consulting Psychology*, 1962, 26, 425-429.
- POLLACK, I. W., & KIEV, A. Spatial orientation and psychotherapy: An experimental study of perception. *Journal of Nervous and Mental Disease*, 1963, 137, 93-97.
- ROSS, W. F. A comparison of personality profiles for campus populations. Unpublished manuscript, University of Kentucky, 1968.
- SANDLER, B. *Investigation of a scale of therapeutic effectiveness: Trust and suspicion in an experimentally induced situation*. (Doctoral dissertation, Duke University) Ann Arbor, Mich.: University Microfilms, 1965. No. 66-1382.
- SHOWS, W. D., & CARSON, R. C. The A-B therapist "type" distinction and spatial orientation: Replication and extension. *Journal of Nervous and Mental Disease*, 1965, 141, 456-462.
- SILVERMAN, J. Personality trait and "perceptual style" studies of the psychotherapists of schizophrenic patients. *Journal of Nervous and Mental Disease*, 1967, 145, 5-17.
- WELCH, R. D., & BERZINS, J. I. The A-B therapist "type" distinction and spatial orientation: Replication. Unpublished manuscript, University of Kentucky, 1968.
- WITKIN, H. A., DYK, R. B., FATERSON, H. F., GOODENOUGH, D. R., & KARP, S. A. *Psychological differentiation: Studies of development*. New York: Wiley, 1962.

(Received October 18, 1968)

DISTINCTIVE PERSONALITY ATTRIBUTES OF CHILD-ABUSING MOTHERS¹

BARRY MELNICK AND JOHN R. HURLEY²

Michigan State University

To explore hypotheses derived from contemporary child-abuse writings, groups of 10 abusive (A) and 10 control (C) mothers, matched for age, social class, and education, were compared on 18 personality variables. As differed reliably from Cs by scoring higher on TAT pathogenicity and dependency frustration, but lower on TAT need to give nurturance, self-esteem (California Test of Personality), Manifest Rejection, and family satisfaction (Family Concept Inventory). These findings are inconsistent with contemporary descriptions of As as being chronically hostile, overwhelmed by maternal responsibilities, domineering participants in a power struggle, or as "normal personalities." Appearing more characteristic of As are an inability to empathize with their children, severely frustrated dependency needs, and a probable history of emotional deprivation.

The recent child-abuse literature offers many hypothetical explanations for this behavior, but few statistically defensible findings. Thus, the child-abuse victim's mother has been variously described as (a) chronically aggressive (Curtis, 1963; Nurse, 1964; Young, 1954), (b) highly frustrated by the restrictions of the maternal role (Blumberg, 1964; Feinstein, 1964; Gladston, 1965), (c) an authoritarian locked into a power struggle with the child (Newston, as quoted in Blumberg, 1964), (d) a psychologically normal individual responding violently to a host of social stresses (Elmer, 1965; Reinhart, 1967), and (e) an unstable, emotionally immature person (Nathan, 1965), who is extremely dependent and narcissistic (Kemp, Silverman, Steele, Droegemueller, & Silver, 1962; Komisaruk, 1966). The role-reversal hypothesis (Morris & Gould, 1963; Steele, Pollock, & Davoren, 1966) holds that the abusive parent invests the child with the negative feelings and critical attitudes earlier held by her own parents, and emphasizes her low self-esteem, dependence, and emotional insecurity. A statistical exploration

of how these ascribed attributes apply to the abusive (A) mother is offered by this study.

METHOD

Ten As referred to the Wayne County Clinic for Child Study by physicians for suspected child-abuse, based on such medical evidence as subdural hematomas, soft tissue swellings, skin bruising, and repeated fractures, were compared with 10 control (C) mothers. Two criteria further defined this A group: (a) the mother, rather than the father, was believed to have inflicted the abuse; and (b) the abused child was under 3 years of age. Probably because this clinic is in a lower-class Negro section of Detroit, eight As were low-socioeconomic-class Negroes. Half of the As resisted returning to the clinic for study and were tested within their homes after being told only that additional information was needed from them. Potential Cs were recruited from the same locale with the aid of the Brewster Neighborhood Social Service Organization of Detroit which supplied a list of young mothers residing within the Brewster Housing Project. The first 10 Cs personally contacted agreed to participate in "a research project concerned with the attitudes and feelings of young mothers"; all were Negro. The following data describe the degree of match achieved between the A and C groups on the selected variables: age in years: $\bar{A} = 23.8$, $\bar{C} = 25.6$; total weekly income: $\bar{A} = \$92.8$, $\bar{C} = \$93.3$; years of education: $\bar{A} = 9.9$, $\bar{C} = 11.6$; number of children: $\bar{A} = 2.5$, $\bar{C} = 2.8$; age when first married: $\bar{A} = 20.3$, $\bar{C} = 20.2$; and Hollingshead-Redlich (1958) social class index³: $\bar{A} = 4.8$, $\bar{C} = 4.8$.

Each S was administered four personality assessment measures selected because of their apparent relevance to certain of the child-abuse hypotheses.

¹ This article is based on the thesis submitted by the first author to the Department of Psychology, Michigan State University. The authors wish to thank Lucy C. Ferguson and Arthur A. Scagull for their helpful suggestions. James Crowder's ratings of the TAT stories are gratefully acknowledged.

² Requests for reprints should be sent to John R. Hurley, Department of Psychology, Michigan State University, East Lansing, Michigan 48823.

³ Because many of the Ss resided in housing projects, the "residence" element of the index was replaced by annual income, divided into seven classes.

Given first were the Self-Reliance, Self-Esteem, Sense of Personal Freedom, Feelings of Belonging, and Withdrawing Tendencies subscales from the California Test of Personality (CTP; Thorpe, Clark, & Tiegs, 1953). The CTP has been validated with various global indexes of personality adjustment (Peak, 1963; Tindall, 1955). Next was the Family Concept Inventory (FCI), a measure of general satisfaction with one's family which has differentiated successfully between families predetermined to be high and low in general adjustment (van der Veen, Huebner, Jorgens, & Neja, 1964) and which also correlates substantially (van der Veen et al., 1964; Palonen, 1966) with the Locke-Wallace Marital Adjustment scale. Then administered was an index of the general harshness of parental disciplinary policies, the Manifest Rejection (MR) scale (Hurley, 1965), shown to correlate importantly with parents' direct, oral acknowledgments that they applied overtly punitive acts to their child.

Last given was a set of 12 TAT cards, the resulting stories being analyzed by two scoring techniques (Melnick, 1968). One approach, called the Pathogenic Index (PI), is scored for content reflecting either an inability to empathize with others or which shows a lack of "basic trust" in the environment. The PI has differentiated as predicted between the mothers of schizophrenics and of nonschizophrenics in several recent studies (Meyer, 1964; Meyer & Karon, 1967; Mitchell, 1965). In the other method TAT stories were scored for needs Affiliation, Aggression, Dependence, Dominance, Independence,

and Nurturance. Need frustration scores were also determined for needs Affiliation, Dependence, Dominance, and Independence; needs Aggression and Nurturance had been defined in ways which made similar scores inappropriate.

RESULTS

All TAT stories were scored independently by the senior author and another rater. The interscorer reliabilities, using the product-moment correlation, exceeded .80 for all need and frustration scores; for PI this $r = .94$.

The t test for small samples (Hays, 1963) was used to assess the significance of all differences between the A and C groups. The small size of the samples and the crudeness of the personality measures were viewed as sufficient justification for using the .10 level of significance in interpreting all results. Table 1 shows that 6 of the 18 measures reliably differentiated between the A and C mothers. The As revealed lower CTP Self-Esteem ($p < .05$), less family satisfaction ($p < .10$), a higher PI ($p < .002$), less need Nurturance ($p < .05$), higher frustration of need De-

TABLE 1
PERSONALITY MEASURE COMPARISONS OF THE ABUSIVE AND CONTROL MOTHERS

Personality measure	M		SD		<i>t</i>
	A	C	A	C	
California Test of Personality					
Self-Reliance	9.2	11.0	3.52	2.36	-1.24
Self-Esteem	8.7	12.0	3.77	1.91	-2.56**
Sense of Personal Freedom	9.5	11.1	3.34	2.18	-1.21
Sense of Belonging	10.5	12.4	4.00	1.22	-1.27
Withdrawing Tendencies	8.8	10.6	4.64	2.45	-1.05
Family Concept Inventory (FCI)	91.4	130.5	49.20	26.90	-2.09*
Manifest Rejection (MR)	54.0	66.0	11.20	10.90	-2.12**
TAT Ratings					
Pathogenic Index (PI)	70.7	39.3	17.24	19.44	3.62***
need Dependence	.27	.23	.18	.14	.59
need Independence	.24	.26	.10	.09	-.24
need Affiliation	.12	.11	.06	.07	.22
need Aggression	.19	.13	.11	.13	1.07
need Dominance	.07	.08	.05	.04	-.42
need Nurturance	.05	.16	.04	.15	-2.36**
n Dependence frustration	.18	.07	.12	.09	2.17**
n Independence frustration	.18	.15	.14	.08	.68
n Affiliation frustration	.12	.09	.08	.08	.66
n Dominance frustration	.06	.05	.05	.07	.31

* $p < .10$, two-tailed.
 ** $p < .05$, two-tailed.
 *** $p < .002$, two-tailed.

TABLE 2

PRODUCT-MOMENT CORRELATIONS AMONG PERSONALITY ATTRIBUTES WHICH DIFFERENTIATED BETWEEN ABUSIVE AND CONTROL MOTHERS

Personality attribute	MR	PI	Self-Esteem	need Nurturance	Dependency frustration
FCI	-.12	-.58*	.78**	.25	-.39
MR		-.24	.09	.26	-.40*
PI			-.43*	-.62**	.25
Self-Esteem				.15	-.32
need Nurturance					.19

* $p < .10$, two-tailed.

** $p < .05$, two-tailed.

pendence ($p < .05$), and a less openly rejectant (MR) stance toward children ($p < .05$).

Table 2 gives the product-moment correlations among these six measures which differentiated between the A and C mothers. Family satisfaction (FCI) relates importantly to self-esteem and inversely to pathogenicity (PI). Self-esteem and an inability to empathize with and accept others (PI) correlated negatively, while this latter variable (PI) also relates inversely with the need to give nurturance.

The intercorrelations among these six measures were subjected to an elementary factor analysis (McQuitty, 1961) which disclosed that the following variable pairs were more highly co-linked than associated with any other index: FCI and CTP Self-Esteem, PI and need Nurturance (need to give nurturance), and MR with need Dependence frustration. In terms of their percentage contributions to the total covariance of this six-variable matrix, these measures ranked: FCI (26%), PI (22%), CTP Self-Esteem (21%), need Nurturance (13%), need Dependence frustration (12%), and Manifest Rejection (7%).

DISCUSSION

Plainly limiting the generalizations which can be drawn from the present findings are the smallness of the samples and their largely lower-class Negro composition. Further, it would have been advantageous to have appraised the degree of disturbance of all the children, or to have had a C group mothers of disturbed children, because the A mothers

might have been confronted with dealing with a more disturbed child than those faced by the C group.

Within these restrictions, the As high PI and low need (to be) Nurturant suggest that these women may have a deficient capacity for empathizing with and administering to their children's needs. These scores also imply that the As had experienced considerable prior frustration of their own emotional needs. Fey's (1955) findings that self-acceptance varies with both the feeling of being accepted by others and the actual acceptance by others seems relevant. The As do not perceive their families as sufficient in meeting their emotional needs. The FCI mean (91.4) of the As was sharply below that of the Cs (130.5). Yet, the Cs mean falls notably below the mean (154.9) of two samples (Palonen, 1966; Updyke, 1968), totaling 99 Ss from a university community of young, middle-class, white mothers. The relationships between the FCI index and the two psychological adjustment indicators (PI and CTP Self-Esteem) suggest that the As psychopathology is associated with and perhaps contributes to their very poor family adjustment.

The As frustrated need Dependence reflects both marked dependency and feelings of being unable to cope with the responsibilities of life accompanied by a perceived lack of environmental emotional supports. This view fits Olive's (1966) data which suggested that child killings often followed shortly after the mother had been abandoned by or cut off from an important supportive person, usually a husband or lover.

It can be speculated that the As may experience a deep sense of anxiety from their feelings of being unable to cope with their responsibilities in the face of inadequate supports which predisposes them to the arousal of anger (Sullivan, 1953, p. 212) whenever their children are difficult to manage. These frustrated dependency needs may lead to a semiobsessive preoccupation with their own needs which limits their ability to identify with the child. Also, this narcissism may hinder the development of a strong attachment to the child which might have helped to control aggressive impulses.

That the As would score below the Cs on Manifest Rejection ($\bar{A} = 54.0$, $\bar{C} = 66.0$) was not anticipated. Defensiveness may be the explanation, for the As could well have feared that psychological test findings might contribute to the initiation of criminal charges. An alternative interpretation is that the As actually had more rigid defenses against the expression of hostility than did the Cs. This view is consistent with Megargee's (1967) report that homicidal killers include both overtly aggressive and overcontrolled types of individuals.

REFERENCES

- BLUMBERG, M. When parents hit out. *Twentieth Century*, 1964, 173, 39-44.
- CURTIS, G. Violence breeds violence—perhaps. *American Journal of Psychiatry*, 1963, 120, 386-387.
- ELMER, E. *The fifty families study*. (NIMH Tech. Rep. 880 (R11)) Washington, D. C.: Government Printing Office, 1965.
- FEINSTEIN, H. Group therapy for mothers with infanticidal impulses. *American Journal of Psychiatry*, 1964, 120, 882-886.
- FEX, W. Acceptance by others and its relation to acceptance of self and others: A reevaluation. *Journal of Abnormal Psychology*, 1955, 50, 274-276.
- GLADSTON, R. Observations on children who have been physically abused and their parents. *American Journal of Psychiatry*, 1965, 122, 440-443.
- HAYS, W. L. *Statistics for psychologists*. New York: Holt, Rinehart & Winston, 1963.
- HOLLINGSHEAD, A. B., & REDLICH, F. C. *Social class and mental illness: A community study*. New York: Wiley, 1958.
- HURLEY, J. R. Parental acceptance-rejection and children's intelligence. *Merrill-Palmer Quarterly*, 1965, 11, 19-31.
- KEMPE, H., SILVERMAN, F., STEELE, B., DROEGE-MUELLER, W., & SILVER, H. The battered-child syndrome. *Journal of the American Medical Association*, 1962, 181, 17-24.
- KOMISARUK, R. Clinical evaluation of families in which child abuse has taken place. Symposium presented at the Workshop on Child Abuse of the Wayne County Juvenile Court, Detroit, May 1969.
- MCQUITT, L. L. Elementary factor analysis. *Psychological Reports*, 1961, 9, 71-78.
- MEGARGEE, E. Matricide, patricide, and the dynamics of aggression. In J. R. Hurley (Chm.), *Murder within the family*. Symposium presented at the meeting of the American Psychological Association, Washington, D. C., September 1967.
- MELNICK, B. Distinctive personality attributes of child-abusing mothers. Unpublished master's thesis, Michigan State University, 1968.
- MEYER, R. A study of the schizophrenogenic mother concept by means of the TAT. Unpublished master's thesis, Michigan State University, 1964.
- MEYER, R., & KARON, B. The schizophrenogenic mother concept and the TAT. *Psychiatry*, 1967, 30, 173-181.
- MITCHELL, K. An elaboration of a study of the schizophrenogenic mother concept by means of the TAT. Unpublished doctoral dissertation, Michigan State University, 1965.
- MORRIS, G., & GOULD, R. W. Role reversal: A concept in dealing with the neglected-battered child syndrome. In *The neglected-battered child syndrome*. New York: Child Welfare League of America, 1963.
- NATHAN, H. Abused children. *American Journal of Psychiatry*, 1965, 122, 443.
- NURSE, S. Familial patterns of parents who abuse their children. *Smith College Studies in Social Work*, 1964, 35, 11-25.
- OLIVE, R. O. Filicide as related to female Oedipal problems. Unpublished doctoral dissertation, Michigan State University, 1966.
- PALONEN, D. J. Interpersonal perceptions and marital adjustment. Unpublished master's thesis, Michigan State University, 1966.
- PEAK, B. California Test of Personality: A study of validation. *Dissertation Abstracts*, 1963, 24, 1281.
- REINHART, J. Child-beating laid to lost ideal. *Newsday*, March 11, 1967.
- STEELE, B., POLLOCK, C., & DAVOREN, E. Observations on the battered child syndrome and the law. Paper presented at the Washington, D. C., Conference on Child Abuse and the Juvenile Court, May 1966.
- SULLIVAN, H. S. *The interpersonal theory of anxiety*. New York: Norton, 1953.
- THORPE, L. P., CLARK, W. W., & TIEGS, E. W. *Manual for the California Test of Personality*. Los Angeles: California Test Bureau, 1953.
- TINDALL, R. Relationship among indices of adjustment status. *Educational and Psychological Measurement*, 1955, 15, 152-160.
- UPDYKE, P. R. Family and role satisfaction among young married women. Unpublished master's thesis, Michigan State University, 1968.
- VAN DER VEEN, F., HUEBNER, B., JORGENSEN, B., & NEJA, P. Relationships between parents' concept of the family and family adjustment. *American Journal of Orthopsychiatry*, 1964, 34, 45-55.
- YOUNG, L. *Wednesday's children: A study of child neglect and abuse*. New York: McGraw-Hill, 1954.

(Received October 21, 1968)

SPI DISCRIMINATORS AMONG FOUR PSYCHOLOGICAL HEALTH-SICKNESS LEVELS¹

DON SPIEGEL

Brentwood Hospital, Veterans Administration Center, Los Angeles

Criterion groups representing four psychological health-sickness levels were formed for each sex from among 87 hospitalized psychiatric patients and 145 normal adults. Health-sickness levels were established for patients on the basis of clinical judgments, and for normals, on the basis of an MMPI severity of pathology index. Magnitudes of mean scores for 14 scales of the Spiegel Personality Inventory (SPI) were in rank order of pathology level for males and females. Both univariate and multivariate methods were used to select 10 scales which appear to be the most promising for use in estimating health-sickness levels in male and female adults. These were deviant response, deterioration feeling, self-dissatisfaction, intrusive thought, alienation, dread, future planlessness, olfactory minimizing, auditory minimizing, and tactual minimizing.

Although a revised system of psychiatric nomenclature has recently been developed in an effort to improve international communication among mental health professionals and in order to meet changing administrative needs and therapeutic goals (Spitzer & Wilson, 1968), considerable work has also been done to provide a basis for describing individual behavior in terms of a single "health-sickness" continuum (Luborsky, 1962). Proponents of the latter approach, which tends to cut across *varieties* of pathology to focus on *levels* of functioning or dysfunctioning, contend that individuals may be meaningfully described and ranked in terms of their overall ability to cope with life situations under ordinary conditions of stress, or by the extent to which adequate psychological functioning is impaired by disease, organic pathology, and/or by maladaptive defenses or behavior patterns.

¹ Contributions of others to this study were as follows: W. J. Dixon (statistical consultation); Arlene Hartry (assistance in project development); Normal Sadwick, Ursula Henderson, Charles Acker, Leonard Rosalsky, Rodolfo Vocal, Carl Hanssen (clinical ratings); Howard Kroll and Joel Abrahams (test scoring); Marie Brady (manuscript typing); and Morris Paulson (assistance in various phases). Data was tabulated by the Statistical Unit, Psychology Service, Brentwood Hospital. Data was transferred to punched cards by Veterans Administration Western Research Support Center.

² Requests for reprints should be sent to Don Spiegel, Brentwood Hospital, Veterans Administration Center, Wilshire and Sawtelle Boulevards, Los Angeles, California 90073.

In recent years, attempts have been made to develop reliable test indexes of "severity of psychopathology," in particular with the MMPI, for example, Tamkin (1959), Sines and Silver (1963), Blumberg (1967), and Dunlop and Waltmann (1966). Also, considerable success has been reported in relating ratings based on Rorschach responses to clinical judgments of health-sickness level (Mayman, 1967).

The present study combined univariate and multivariate approaches to the problem of selecting sets of test variables from the Spiegel Personality Inventory (SPI; Spiegel, 1965) which would maximize the likelihood of accurately classifying members of criterion groups representing four levels of psychological health-sickness and would minimize the likelihood of misclassification.

METHOD

Subjects

All patients on a male ward and a female ward of a Veterans Administration neuropsychiatric hospital were rated for severity of psychopathology on a 21-point scale on which health-sickness levels were explicitly defined.² Each patient was rated by his

² The rating scale was subdivided as follows: 1-3 points (ideal normal), 4-9 (usual range for non-hospitalized "normals"), 10-18 (usual range for testable psychiatric patients), and 19-21 (usually too severely ill to be tested). Each range of 3 points was further defined. Raters were instructed that most ratings should fall in the range from 10 to 18 points which is a range of moderate to markedly severe psychopathology characterized by increasing

ward psychologist and psychiatrist. A battery of tests was administered to the patients in small supervised sessions. The criteria for including a patient in the study were (a) the two pathology ratings had to differ by less than six points, (b) all tests had to be completed and show evidence that *S* understood and cooperated with instructions, and (c) MMPI validity indexes had to show no evidence that an attempt was made to present an overly-favorable self-image on the test. The ratings of pathology made by psychologist-psychiatrist pairs were added. For each sex, two pathology levels (i.e., a "sicker" level (D) and a "healthier" level (C)) were formed on the basis of combined ratings. Twenty-one male patients and 23 female patients who received combined ratings equal to or less than the median combined rating of 26 were at the healthier level. Twenty-one male patients and 22 female patients whose combined ratings were above 26 were at the sicker level.

Normal *Ss* were drawn from several sources in an attempt to match the patient groups as closely as possible on variables including age, socioeconomic level, education, and general occupational level. Most *Ss* were between 30 and 50 years of age, Caucasian, and from the lower to lower-middle socioeconomic level. Normals were similarly tested and subdivided into a level without apparent pathology (A) and a level with apparent "mild" pathology (B) on the basis of the Multiphasic Index (*MI* scale) of the MMPI. This scale, based on a weighted configuration of other scale scores, was reported by Dunlop and Waltmann (1966) to be a useful index of severity of psychopathology. The mean *MI* score for the 145 male and female normals (A plus B groups) was 87.45. Level A was represented by 44 males and 25 females with *MI* scores of 87 and below, and Level B was represented by 29 males and 47 females with scores above 87.

Since the normals and patients were necessarily classified on the basis of different criteria of health-sickness, a further check was made on the validity of the assumption that group order ABCD formed a rank order of health-sickness levels. The mean scores and standard deviations of 20 MMPI clinical and research scales (*MI* excluded) for each group were given to six clinical psychologists to rank independently for severity of psychopathology. There was unanimous agreement that ABCD was the proper ordering of groups on a health-sickness dimension.

Tests and Scales

The test battery included the complete MMPI and SPI, in addition to nine other tests (not to be considered here). Description of special MMPI scales may be found in Dahlstrom and Welsh (1960). Development of the SPI has been described elsewhere in terms of difficulty in satisfactorily coping with problems and functioning adequately under stress. Raters were also asked to indicate their level of confidence in rating each patient (high, moderate, or low).

where (Elashoff & Spiegel, 1969; Spiegel & Hartry, 1968; Spiegel & Keith-Spiegel, 1969), and a manual is available for research use (Spiegel, 1965). For the present study the complete 200-item inventory was administered and 41 scales scored and used in the analyses. Thirty-five of these scales are either 15- or 18-point Guttman scales with no item overlap (Guttman, 1950). Of particular interest in this research was a special deviant response scale composed of 39 items from these scales which normals and psychiatric patients had answered differently on a previous study (with chi-square values for each item significant at $p < .01$), a two-item deterioration feeling scale (which calls for admission or denial of regression in one's level of functioning), and a set of sensory and environmental "maximizer-minimizer" scales (which measure preference for relatively high, moderate, or low levels of stimulation). Scale names reflect the high end of the scoring range which, in most instances, is the pathological end of a psychological dimension.

RESULTS

Table 1 gives the means and *F* values derived from univariate analyses of variance of MMPI scales on 21 variables for each sex. As may be seen, the mean MMPI profiles of Level A men and women were similar and appear quite healthy. Level B men and women also had similar profiles, but with an elevation above Level A profiles on most scales which suggests the presence of mild emotional disturbance. Males tended to be somewhat higher on *Hy*, *Pd*, and *Ma* and females slightly higher on *A_T* and *Dy*. MMPI mean scores showed considerably greater separation between B and C levels of health-sickness for men than for women, which probably accounts for the much larger *F* values for males than for females. On most scales both C and D level females showed less pathology than either C or D level males (e.g., note mean scores on *Hs*, *D*, *Hy*, *Pd*, *Pa*, *Pt*, *Sc*, *Ma*, *A*, *R*, *Es*, *A_T*, and *MI*). The D group means (both male and female) showed more psychotic and sociopathic symptomatology than the C group means which, in turn, revealed higher neurotic pathology than the D group means. Thus, although ABCD was apparently the appropriate rank order of mean pathology for each sex, male patients with a given rating of pathology tended to have higher MMPI scores than females with the same rating. In view of this, it appeared advisable that separate discriminant analyses should be done on the data from males and females.

TABLE 1

MMPI MEANS AND *F* VALUES FOR MALE AND FEMALE ADULTS OF FOUR HEALTH-SICKNESS LEVELS

MMPI scale	Level—Males				<i>F</i>	Level—Females				<i>F</i>
	A	B	C	D		A	B	C	D	
<i>Hs</i>	48.7	52.3	69.2	62.9	16.36**	49.6	51.4	57.0	56.2	3.30**
<i>D</i>	51.0	56.0	79.9	78.0	38.89**	49.8	56.0	70.8	71.8	22.43**
<i>Hy</i>	55.9	57.0	70.5	62.0	9.55**	54.4	52.2	59.4	57.9	3.32*
<i>Pd</i>	54.5	61.8	71.3	72.7	15.27**	56.6	54.0	64.0	70.4	10.76**
<i>Mf</i>	56.8	55.2	66.1	63.3	8.02**	48.4	49.7	47.1	48.6	.37
<i>Pa</i>	51.0	55.4	67.9	78.1	30.98**	52.7	56.4	62.0	67.2	8.15**
<i>Pt</i>	48.8	56.5	76.0	78.1	41.68**	50.2	55.0	65.0	71.0	22.71**
<i>Sc</i>	48.4	56.5	80.8	92.3	48.54**	51.0	55.4	68.3	78.0	23.32**
<i>Ma</i>	56.2	61.1	61.6	65.8	3.73	50.7	53.7	58.0	58.0	2.40
<i>Si</i>	45.8	49.7	58.6	64.1	25.77**	49.4	59.4	61.9	64.8	15.86**
<i>L</i>	3.6	3.6	4.2	4.5	1.13	4.6	3.4	4.4	4.4	3.73*
<i>F</i>	2.6	5.8	10.4	17.8	30.63**	2.2	5.4	8.8	11.8	14.30**
<i>K</i>	16.4	14.1	14.0	12.3	5.39**	17.4	12.7	12.3	12.5	8.57**
<i>AI</i>	47.1	58.2	89.4	98.9	37.24**	47.3	60.6	81.7	90.5	28.90**
<i>A</i>	41.7	49.9	60.2	64.9	42.00**	40.2	50.9	56.8	59.3	20.39**
<i>R</i>	48.5	45.7	54.9	53.4	5.37**	47.2	48.6	51.6	51.2	.82
<i>Es</i>	62.6	54.4	40.9	34.1	66.31**	59.5	52.5	44.6	41.9	13.90**
<i>F/NF</i>	1.1	3.3	8.2	13.5	39.98**	1.1	3.7	6.3	8.5	14.10**
<i>A_T</i>	9.3	14.7	24.9	24.9	40.50**	10.9	18.6	23.5	24.1	17.10**
<i>Dy</i>	12.9	19.0	28.8	33.4	44.15**	16.2	27.5	31.3	33.3	20.91**
<i>MI</i>	80.0	92.8	112.0	115.8	57.37**	80.8	94.1	105.9	111.0	24.75**

* $p < .05$.** $p < .01$.

Means and standard deviations were computed for all SPI scales and separate analyses of variance were done for each scale to test the equality of means for Groups A, B, C, and D for each sex. These were examined to find scales with mean scores in ABCD order of magnitude for both sexes. The 14 such scales found are among the 18 scales shown in Table 2. Of the variables listed, 15 univariate *F* values were significant at $p < .01$.

Multiple discriminant analyses were then done for sexes separately for the four groups (A, B, C, and D) using the 18 variables shown in Table 2. The analysis (Dixon, 1967) proceeds in a stepwise manner by forming linear sums of first one, then two, then three, etc., variables. The variable added is always the one which gives the greatest improvement in the classification of cases and the greatest decrease in the ratio of within to total generalized variance.

For the discriminant analyses, any variable whose *F* value to enter a set of discriminants was below 2.20 was excluded from entry. For males, three variables were included in

the discriminating set before the *F* values became insufficient for further computation. The single best variable was deviant response which, by itself, misclassified only 8% of the normal males as patients and less than 12% of the patients as normals. The alienation scale improved the accuracy of classifying members of Subgroups C and D, and intrusive thought increased the accuracy of classifying both A and D groups. The *F* values for all between-groups comparisons were significant at $p < .01$. The percentages of cases classified into Groups ABCD by these discriminants are shown in Tables 3. This set of three variables correctly classified 77% of A group males, 39% of B group, 48% of C group, and 62% of D group.

For females, five variables were included in the discriminating set before the *F* values became insufficient for further computation. Again, the best single variable was deviant response which did not do as well in classifying females as it did males. In fact, it misclassified 25% of the normal females as patients and 35% of the patients as normals. The

TABLE 2

SPI MEANS AND *F* VALUES FOR MALE AND FEMALE ADULTS OF FOUR HEALTH-SICKNESS LEVELS

SPI scale	Level—Males				<i>F</i>	Level—Females				<i>F</i>
	A	B	C	D		A	B	C	D	
Deviant response	66.7	75.3	104.7	106.1	53.80**	76.1	86.2	104.0	105.4	21.91**
Deterioration feeling	2.9	3.2	6.0	5.9	30.99**	3.4	3.9	5.5	6.3	13.80**
Self-dissatisfaction	9.5	12.0	16.2	15.9	23.27**	10.6	12.8	16.4	16.6	11.65**
Social dissatisfaction	10.6	12.1	15.1	16.4	16.57**	10.4	13.7	16.7	17.2	19.60**
Social insecurity	8.0	10.9	12.6	13.1	14.71**	8.4	11.2	13.1	13.5	10.34**
Alienation	7.7	8.6	10.1	13.4	19.63**	7.8	9.8	10.8	11.5	7.09**
Fearfulness	9.3	9.7	11.6	12.1	5.34**	12.4	13.0	13.5	14.7	1.92
Dread	8.0	9.5	12.4	13.8	18.55**	8.0	9.8	11.0	13.1	8.61**
Hopelessness	7.4	6.9	11.0	11.7	3.93*	6.1	7.1	9.9	11.0	16.32**
Anger-depression	10.3	12.4	15.5	16.0	12.03**	12.8	13.1	15.0	15.5	2.72*
Sex disturbance	23.4	27.0	31.3	31.6	12.13**	20.4	25.1	27.7	28.1	9.31**
Impulsive action	10.1	11.9	13.0	13.8	6.32**	9.3	11.9	12.7	12.5	4.32**
Intrusive thought	9.9	13.1	16.9	17.5	30.23**	10.9	13.6	17.2	18.4	14.95**
Obsessiveness	10.4	11.7	14.7	14.9	10.14**	10.6	12.3	14.3	15.2	7.04**
Future planlessness	10.5	10.8	12.4	14.2	5.25**	9.7	11.0	11.5	13.6	4.98**
Olfactory minimizing	8.2	10.1	11.7	12.5	9.36**	9.9	10.5	10.7	13.8	5.70**
Auditory minimizing	8.2	10.1	13.8	14.5	26.89**	9.6	11.4	12.1	15.3	11.10**
Tactual minimizing	7.6	10.0	11.3	13.6	18.76**	10.0	10.1	12.3	14.2	9.94**

Note.—*F* values are based on univariate analyses of variance.* $p < .05$.** $p < .01$.

addition of social dissatisfaction improved the classification of normals into Subgroups A and B; and the addition of auditory minimizing improved the classification of C and D patients. Tactual minimizing and deterioration feeling improved the classification of A and B normals and B and C patients. The *F* values for all between-groups comparisons for the five-variable set were significant at $p < .01$. Table 3 shows the percentage of cases classified into Groups ABCD. This set correctly classified 64% of Group A normals, 45% of B normals, 57% of C patients, and 55% of D patients. It misclassified about 21% of the normals (A plus B) as patients (C plus D), and almost 22% of the patients as normals.

The first set of discriminants was then removed from the pool of potential entering variables, and for each sex, a second discriminant analysis was done to determine whether a second linear combination of variables could be found which might do as well or better than the first in classifying cases. For males, four variables were entered into the discriminating set in the order: deteriora-

tion feeling, tactual minimizing, self-dissatisfaction, and auditory minimizing. This set of variables correctly classified cases for male Ss as follows: A: 73%; B: 52%; C: 43%; and D: 53%. Misclassified as patients were 2% of A and 20% of B group normal males; misclassified as normals were 24% of C and 14% of D group male patients. For females, five variables were entered in

TABLE 3

PERCENTAGE OF CASES CLASSIFIED INTO GROUPS BY DISCRIMINATING VARIABLES

Actual group	Assigned group			
	A	B	C	D
Males				
A	77	21	0	2
B	38	39	17	6
C	5	14	48	33
D	0	5	33	62
Females				
A	64	28	4	4
B	27	45	19	9
C	4	17	57	22
D	4	18	23	55

TABLE 4
CORRELATIONS BETWEEN SPI AND MMPI SEVERITY OF PATHOLOGY SCALES

SPI scale	MMPI Scale						
	<i>F</i>	<i>MI</i>	<i>A</i>	<i>A_T</i>	<i>Es</i>	<i>Sc</i>	<i>D-O</i>
Deviant response	61	75	74	75	-75	69	80
Deterioration feeling	42	60	62	62	-60	55	66
Alienation	51	54	50	45	-50	52	56
Dread	51	60	66	62	-63	55	56
Future planlessness	30	26	35	25	-40	29	26
Intrusive thought	48	65	75	74	-70	57	67
Tactual minimizing	40	48	43	44	-53	46	46
Auditory minimizing	53	59	57	57	-61	55	58
Olfactory minimizing	35	31	35	32	-39	34	31
Self-dissatisfaction	38	64	62	65	-49	50	70

Note.—A correlation of .17 is significant at $p < .01$.

the following order: hopelessness, intrusive thought, self-dissatisfaction, and sex disturbance. This set of variables correctly classified cases for female Ss as follows: A: 80%; B: 45%; C: 39%; and D: 50%. Misclassified as patients (either C or D) were 4% of A and 21% of B group normal females; misclassified as normals were 39% of C and 14% of D group female patients.

Thus, from the two discriminant analyses of four groups for each sex, it was found that six variables appeared for both sexes in either the first or second set of discriminants, namely, deviant response, deterioration feeling, self-dissatisfaction, intrusive thought, auditory minimizing, and tactual minimizing. The classification of males into health-sickness levels was aided by alienation, and the classification of females by social dissatisfaction, hopelessness, and sex disturbance.

Pearson product-moment coefficients of correlation were computed between each of the 18 SPI variables shown in Table 1 and clinical ratings of health-sickness level for the total group of 87 male and female patients. The following significant correlations were obtained: alienation (.29), dread (.30), hopelessness (.27), intrusive thought (.29), future planlessness (.28), auditory minimizing (.32), olfactory minimizing (.24), and tactual minimizing (.25). The first four correlations were significant at $p < .01$; the last two at $p < .05$.

The interrelationships between some of the

more promising SPI indexes of health-sickness level and some of the more established MMPI counterparts may be seen in Table 4. These Pearson correlation values are based on the entire group of 232 patients and normals of both sexes. As may be seen, the deviant response scale has extremely high correlations with all the MMPI indexes, with those of deterioration feeling, dread, intrusive thought, and auditory minimizing being close behind. The highest correlation values for deviant response, deterioration feeling, and self-dissatisfaction are with D-Obvious. Future planlessness, tactual minimizing, auditory minimizing, and olfactory minimizing have their highest relationships (negative) with Ego-Strength (*Es*). Intrusive thought is most strongly related to indexes of anxiety (*A*, *A_T*). The scales with weakest relationships to MMPI scales are future planlessness and olfactory minimizing (both related more to *Es* than to the other scales). In a factor analysis of 65 variables including many MMPI and SPI scales, and involving the same 232 cases, the variables in Table 4 all had substantial loadings on a severity of psychopathology factor, with especially high loadings for *A*, deviant response, *MI*, *Es*, intrusive thought, dread, *F*, deterioration feeling, and self-dissatisfaction, in descending order of magnitude.

DISCUSSION

The finding that mean scores on 14 SPI scales were in rank order of pathology level

as defined by criterion groups of male and female adults is of particular importance for further research with the SPI in this area. Examination of Table 1 indicates that good potential discriminators also existed for the MMPI, namely, *Pd*, *Pa*, *Pt*, *Sc*, *Si*, *F*, *AI*, *A*, *Es*, *F/NF*, *Dy*, and *MI*. Indeed, a number of these have been reported elsewhere to be good indexes of severity of pathology.

The 18 SPI scales in Table 2 show good potential for identifying the health-sickness level of respondents. Perhaps the most promising scales are those which had means in ABCD order of magnitude, which had *F* values significant at $p < .01$ in the univariate analyses of variance, and for which scores for patients correlated significantly with criterion ratings of pathology. These, for both sexes, were alienation, dread, intrusive thought, future planlessness, olfactory minimizing, auditory minimizing, and tactual minimizing. Although deviant response, deterioration feeling, and self-dissatisfaction were selected in the discriminant analyses, primarily on the basis of their separation of A from B and B from C groups, they did not appear to be particularly useful in separating healthier from sicker patients.

A novel aspect of the results was the finding that sensory scales of the SPI appear to be among the promising discriminators of health-sickness level. These scales all involve either a preference for or aversion to relatively intense, varied, or abrupt stimulation in a particular modality. Persons earning high scores acknowledge strong aversive reactions to stimulus inputs which low scorers find pleasant or, at least, not annoying. In another study, significant correlations were found between SPI sensory minimizing scores and anxiety scores on the Taylor scale (*A_T*) for both patients and normals. It was also found in previous work, that sensory minimizers selected lower levels of stimulus intensity than did sensory maximizers in free choice laboratory situations. Nevertheless, the significance of sensory minimizing tendencies for psychopathology is unclear at this time.

It has been frequently reported that "sicker" individuals tend to endorse more socially undesirable responses or admit more pathology than healthier individuals. Whether,

in a given instance, this is associated with a lack of effective defenses, a lack of concern for what others think about one, a desire to look ill, or the presence of truly greater pathology, is not always clear. On the basis of the author's analysis of responses to the MMPI and SPI, it would appear that the average person at each successively higher pathology level tends to complain more than those at lower levels about something being wrong, either with himself, or with things going on in the world around him.

Although many SPI items have neither socially desirable nor undesirable response possibilities, it is not these items which appear to be the best discriminators among criterion pathology levels. Even the sensory scales, in which the items do not provide an opportunity to admit or deny pathology per se, do permit *S* to admit annoyance or aversion to certain kinds of stimulus inputs. It would seem that given the opportunity to admit annoyance, the "sicker" individuals will do so. This view was supported by other results obtained from *Ss* used in this study who were asked to indicate the presence of personal problems from comprehensive checklists. The mean number of problems checked were in ABCD order of magnitude for both sexes, and the *F* values for the analyses of variance of four groups were significant at $p < .01$. Thus, it would appear likely that severity of psychopathology or health-sickness level is positively associated with general annoyance level which may be manifested in any suitable vehicle which offers an opportunity to complain.

In terms of the SPI scales of Table 2, which were grouped according to spheres of content similarity, a characterization of the average patient at Brentwood Hospital, male or female, who is regarded by mental health professionals to be among the most severely ill is provided. He is dissatisfied with himself and with social relationships. He feels he is held in disrespect and low regard by others, and he feels misunderstood and isolated. He has a foreboding anticipation of imminent catastrophe, fear that participation in potentially dangerous life events will bring him harm, or a fear of death or attack. He views life as a rather hopeless struggle, with a future which has nothing to offer. He sees himself as easily

upset and in a bad mood much of the time. He feels ill at ease around others and often guilty without knowing why. He tends to fly off the handle easily and to be unable to delay gratification. He has a disturbing inability to maintain control of thoughts, to concentrate and attend. He denies concern with the future and seems unable to plan ahead. He is annoyed by intense or highly varied stimulation in one or more sense modalities.

It is, of course, not to be expected that all individuals at a given pathology level will have similar elevations on the SPI scales. Rather, as with the MMPI, configurations of these and other SPI scale scores tend to elucidate the predominant characteristics of the individual. It is probable that a patient with a generally elevated profile or high elevations on a number of the 18 scales of Table 2, would be regarded as relatively ill, maladjusted, disturbed, or sick by the hospital staff, using criteria other than tests on which to base their judgments.

REFERENCES

- BLUMBERG, S. MMPI *F* scale as an indicator of severity of psychopathology. *Journal of Clinical Psychology*, 1967, 23, 96-99.
- DAHLSTROM, W. G., & WELSH, G. S. *An MMPI handbook: A guide to use in clinical practice and research*. Minneapolis: University of Minnesota Press, 1960.
- DIXON, W. J. *BMD biomedical computer programs*. Berkeley: University of California Press, 1967.
- DUNLOP, E., & WALTSMANN, R. H. The emotional temperature. Paper presented at the First International Congress of the Academy of Psychosomatic Medicine, Palma de Mallorca, Spain, September 1966.
- ELASHOFF, J. D., & SPIEGEL, D. E. Optimal choice of rater teams, II: Applications. *Psychometrika*, 1969, 34, 33-44.
- GUTTMAN, L. The basis for scalogram analysis. In S. A. Stouffer, L. Guttman, E. Suchman, P. Lazarsfeld, S. Star, & G. Clausen (Eds.), *Measurement and prediction*. Princeton, N. J.: Princeton University Press, 1950.
- LUBORSKY, L. Clinicians' judgments of mental health. *Archives of General Psychiatry*, 1962, 7, 407-417.
- MAYMAN, M. Object-representations and object-relationships in Rorschach responses. *Journal of Projective Techniques and Personality Assessment*, 1967, 31, 17-24.
- SINES, L., & SILVER, R. An index of psychopathology (Ip) derived from clinicians' judgments of MMPI profiles. *Journal of Clinical Psychology*, 1963, 19, 324-326.
- SPIEGEL, D. E. *The Spiegel Personality Inventory manual*. Los Angeles: Author, 1965.
- SPIEGEL, D., & HARTY, A. The prediction of severity of emotional disturbance. Paper presented at the meeting of the Western Psychological Association, San Diego, March 1968.
- SPIEGEL, D., & KEITH-SPIEGEL, P. Factor analysis of 78 variables from nine personality tests and scales. *Journal of Projective Techniques and Personality Assessment*, 1969, 33, 160-167.
- SPITZER, R. L., & WILSON, P. T. A guide to the American Psychiatric Association's new diagnostic nomenclature. *American Journal of Psychiatry*, 1968, 124, 1619-1629.
- TAMKIN, A. An MMPI scale measuring severity of psychopathology. *Journal of Clinical Psychology*, 1959, 15, 16.

(Received October 21, 1968)

PROSPECTIVE SPAN AS A COGNITIVE ABILITY

LOUIS S. DICKSTEIN¹

City College of the City University of New York

Prospective span (PS) in TAT stories may be regarded as a measure of the cognitive capacity of anticipation. As such, it should be related to cognitive tasks involving the specific capacity of anticipation but not to more general measures of verbal intelligence. In a sample of 51 male college students, PS was found to be related to two cognitive tasks requiring anticipation, the Mazes Test and the Seeing Deficiencies Test, but unrelated to the Vocabulary and Information subtests of the WAIS. Contrary to previous studies, PS was found to be negatively related to academic grades. Large differences were obtained between PS scores for the different TAT cards.

Within the past 15 years, a number of studies of future time perspective have been conducted which have been variously referred to as studies of foresight, anticipation, planning ability, or prospective span (PS). One of the most promising methods for the measurement of this variable has involved the scoring of TAT stories for PS (Epley & Ricks, 1963; Teahan, 1958).

Epley and Ricks (1963) reported a correlation of only .08 between PS and the Scholastic Aptitude Test. Similarly, in a subsequent study (Ricks, Umbarger, & Mack, 1964), a nonsignificant correlation was obtained between PS and intelligence. On the other hand, Epley and Ricks did report a significant relationship between PS and Cattell's intelligence scale measuring "the capacity to deduce relationships," while Dickstein and Blatt (1967) found that high and low scorers on the Picture Arrangement subtest of the Wechsler Adult Intelligence Scale (WAIS) differed significantly on PS. A possible resolution of these discrepancies may lie in viewing PS as a specific cognitive ability which is significantly correlated with other measures of anticipation, such as the Picture Arrangement subtest, but which is not necessarily related to more general measures of verbal intelligence. The present study was conducted to determine the relationships between PS and measures of general verbal intelligence and measures of the more specific cognitive ability of anticipation.

¹ Requests for reprints should be sent to Louis S. Dickstein, who is now at the Department of Psychology, Wellesley College, Wellesley, Massachusetts 02181.

Prospective span has been found to be related to academic achievement. In a study by Teahan (1958), high and low achievers in school differed significantly in PS, while Epley and Ricks reported a correlation of .43 between PS and 4-year college grade averages. A secondary aim of the present study was to replicate the reported relationship between PS and academic achievement.

METHOD

The Ss were 51 male college students at the City College of the City University of New York. All Ss were volunteers from introductory psychology sections. There were 15 freshmen, 35 sophomores, and 1 junior. Almost all Ss were 18 or 19 years of age.

All Ss were given six cards of the TAT and were required to write stories. The cards used were 1, 2, 8BM, 6BM, 5, and 12M in that order. The stories were scored for PS according to the scoring system developed by Epley and Ricks (1963). In addition, Ss were given four tests. These were the Vocabulary and Information subtests of the WAIS (Wechsler, 1955) and the Maze Tracing Speed Test (MTST) and the Seeing Deficiencies Test (SDT) from the Kit of Reference Tests for Cognitive Factors (French, Ekstrom, & Price, 1963). The Vocabulary and Information subtests were selected as the two verbal subtests with the highest reliability coefficients and with the highest correlations with total verbal IQ (Wechsler, 1955). The raw scores on these subtests were converted to scaled scores. The other two tests were selected because they require anticipation for successful performance. The MTST requires S to find and mark an open path through a moderately complex series of paper mazes. The SDT requires S to point out the way in which a described plan or activity is faulty.

The grade-point index was available for 45 of the 51 Ss and these data were used to check the relationship between grades and PS.

TABLE 1

MEANS, STANDARD DEVIATIONS, AND RELIABILITY
COEFFICIENTS FOR PS AND FOUR
COGNITIVE MEASURES

Measure	<i>M</i>	<i>SD</i>	Reliability
Prospective Span	2.74	1.63	.574
Maze Tracing	22.78	6.87	.921
Seeing Deficiencies	7.75	2.86	.588
Vocabulary	12.92	1.39	.940*
Information	14.02	2.38	.910*

* The reliabilities for the Vocabulary and Information subtests are the reliabilities reported in the WAIS manual (Wechsler, 1955).

RESULTS

The means, standard deviations, and reliability coefficients for the five measures are presented in Table 1. The reliability coefficients for PS, MTST, and SDT are split-half coefficients corrected for full length by the Spearman-Brown formula. It should be noted that although both the MTST and the SDT are speed tests, they are divided into two sections which are timed and scored separately so that a split-half coefficient is appropriate (Helmstadter, 1964). The reliability coefficients for the Vocabulary and Information subtests are those reported in the WAIS manual (Wechsler, 1955). They are also split-half coefficients corrected by the Spearman-Brown formula.

The reliabilities of the Vocabulary, Information, and MTST tests are all excellent, but reliabilities of the PS score and the SDT leave much to be desired. The TAT reliability of .574 is considerably lower than the .80 test-retest reliability which has been reported

TABLE 2

CORRELATION MATRIX OF PS, FOUR COGNITIVE
MEASURES, AND ACADEMIC INDEX

Measure	1	2	3	4	5	6
1. PS	—					
2. MTST	.293*	—				
3. SDT	.270*	.417**	—			
4. Vocabulary	-.018	.289*	.202	—		
5. Information	.138	.425**	.374**	.645**	—	
6. Index	-.221	.012	.088	.090	-.014	—

Note.—*N* = 51 for the correlations between the cognitive measures and 45 for all correlations with the Index.

* $p < .05$, one-tailed.

** $p < .01$, one-tailed.

(Ricks et al., 1964). Furthermore, the average intercorrelation between the six stories in the present sample is only .224 which is not significantly greater than chance. Nine of the 15 intercorrelations, however, were significant with the highest reaching .372. Forty of the Ss' protocols were also scored by a second rater. The product-moment interrater reliability for PS was .959.

The correlation matrix for the five measures plus academic index is presented in Table 2. As expected, the PS score is significantly correlated with both the MTST and the SDT, although the magnitude of the correlations is modest. The MTST and the SDT are also significantly correlated with each other. The multiple correlation coefficient between PS and the MTST and the SDT is .335 ($p < .01$, one-tailed). It is instructive to note that while the MTST and SDT are both significantly correlated with the Information subtest, and the MTST is also significantly correlated with the Vocabulary subtest, the correlations between PS and the WAIS subtests are not significant.

As previously noted, the reliabilities of the PS score and the SDT are rather low. If the correlations are corrected for attenuation, the correlation between PS and the MTST becomes .403, between PS and the SDT, .465, and between the MTST and the SDT, .567 (For all of these, $p < .01$, two-tailed). The correlations between PS and the WAIS subtests remain nonsignificant.

The correlations between the grade-point index and the various measures are somewhat surprising. The index is not significantly related to any of the measures including the WAIS subtests. The highest correlation is with the PS score and that correlation is the

TABLE 3

MEANS, *t* VALUES, AND PROBABILITY LEVELS FOR
COMPARISON OF HIGH- AND LOW-PS SUBJECTS

Measure	High	Low	<i>df</i>	<i>t</i>	<i>p</i> ^a
MTST	26.18	22.34	32	1.70	.05
SDT	8.18	6.59	32	1.87	.05
Vocabulary	13.12	13.12	32	.00	<i>ns</i>
Information	14.53	14.18	32	.53	<i>ns</i>
Index	1.98	2.36	28	2.26	.025

^a One-tailed.

reverse of what might be expected from previous studies. Indeed, when this correlation is corrected for the low reliability of the PS score, the correlation becomes $-.292$ ($p < .10$, two-tailed).

In a further analysis, the top third and bottom third of Ss on PS were selected and t tests were conducted for the differences between these groups on all measures. The mean PS score for the 17 high-PS Ss was 4.58, while the mean for the 17 low-PS Ss was 1.00. The means, t values, and probability levels for the five comparisons are presented in Table 3. The differences between the groups are significant for the MTST, the SDT, and academic index. The differences between the groups on academic index is the most significant. The groups are not significantly different on the Vocabulary and Information subtests.

Finally, it should be noted that the different TAT cards elicit different prospective spans. The mean PS for the six TAT cards in terms of their order of presentation were 2.16, 3.49, 3.61, 3.45, 1.84, and 1.94. This point has been noted before by Ricks (1965).

DISCUSSION

The results of the study support the first hypothesis. The PS score on the TAT is significantly related to other measures of anticipation but not to general verbal intelligence. Admittedly, the relationship is a moderate one which, considering the multiple correlation, accounts for only about 11% of the variance of PS. On the positive side, it does not seem possible to attribute the obtained relationships to common method variance as is the case for so many personality variables (Mischel, 1968), since the three measures entail quite divergent operations. Rather, the correlations appear to represent a meaningful relationship which would probably be considerably higher were the methods of measurement more reliable.

The low reliability of the PS score may be attributed, in part, to the use of only six cards in the present study. A minimum of 10 cards would appear to be a more appropriate procedure. As for the SDT, the extremely low mean of 7.75 out of a possible score of 20 for a college population suggests

that the time limit of 10 minutes per section is too stringent. A longer time period would allow more time for thought and probably would also result in higher reliability for that measure. In any event, the goal of developing an adequate instrument for the measurement of such an important capacity as anticipation would seem to be highly worthwhile.

The absence of a significant correlation between PS and the Information and Vocabulary subtest scaled scores is limited to the range of IQ in the sample. This range may be estimated from the subtest scaled scores as between 97 and 141, although 84% of the scores fall between 112 and 138. The mean estimated IQ is 123. Although this is a restricted range, it is sufficiently broad to give meaning to the absence of any relationship.

The differences between mean PS for the different TAT cards suggests the need for more systematic study of the role of stimulus variables. It certainly seems inappropriate to use any set of TAT cards and to assume that the results are comparable to studies based on other sets of cards.

The negative relationship between PS and academic index was completely unexpected and contradicts two previous studies. No explanation can be offered for this finding at the present time.

Finally, it should be noted that all studies of PS thus far have used male samples. Studies of anticipation in females would seem to be in order.

REFERENCES

- DICKSTEIN, L. S., & BLATT, S. J. The WAIS Picture Arrangement subtest as a measure of anticipation. *Journal of Projective Techniques*, 1967, 31, 32-38.
- EPLEY, D., & RICKS, D. Foresight and hindsight in the TAT. *Journal of Projective Techniques*, 1963, 27, 51-59.
- FRENCH, J. W., EKSTROM, R. B., & PRICE, L. A. *Manual for kit of reference tests for cognitive factors*. Princeton: Educational Testing Service, 1963.
- HELMSTADTER, G. C. *Principles of psychological measurement*. New York: Appleton-Century-Crofts, 1964.
- MISCHEL, W. *Personality and assessment*. New York: Wiley, 1968.
- RICKS, D. Measurement of temporal orientation and span through the use of fantasy materials. Paper presented at the meeting of the Eastern

- Psychological Association, Atlantic City, April 1965.
- RICKS, D., UMBARGER, C., & MACK, R. A measure of increased temporal perspective in successfully treated adolescent delinquent boys. *Journal of Abnormal and Social Psychology*, 1964, 69, 685-689.
- TEAHAN, J. Future time perspective, optimism, and academic achievement. *Journal of Abnormal and Social Psychology*, 1958, 57, 379-380.
- WECHSLER, D. *WAIS manual*. New York: The Psychological Corporation, 1955.

(Received October 25, 1968)

(Continued from page 738)

- A Dimension of Moral Judgment: Robert Hogan*: Department of Psychology, Johns Hopkins University, Baltimore, Maryland 21218.
- Treatment of "Adolescent Rebellion" by the Reconstual of Stimuli: John C. Masters*: Institute of Child Development, University of Minnesota, Minneapolis, Minnesota 55455.
- College Students as Companions the Long-Term Mental Hospital Patients: Some Considerations: Edward H. Fischer*: Department of Psychology, Connecticut Valley Hospital, Middletown, Connecticut 06457.
- Autonomic Reactivity to Sensory Stimulation in Psychopatic, Neurotic, and Normal Juvenile Delinquents: Thomas D. Borkovec*: Psychological Clinic, Childrens Research Center, Champaign, Illinois.
- Anxiety and Speech in the Initial Interview: Benjamin Pope,* Aron W. Siegman, and Thomas Blass: Department of Psychology Service, Sheppard and Enoch Pratt Hospital, Baltimore, Maryland 21204.
- Approval Motive and Academic Behaviors: The Self-Reinforcement Hypothesis: Michael S. Matell and Ronald E. Smith*: Department of Psychology, University of Washington, Seattle, Washington 98105.
- Anxiety and Depression in Speech: Benjamin Pope,* Thomas Blass, Aron W. Siegman, and Jack Rahe: Department of Psychology Service, Sheppard and Enoch Pratt Hospital, Baltimore Maryland 21204.
- A Semantic Space for Nonverbal Behavior: Albert Mehrabian*: Department of Psychology, University of California, Los Angeles, California 90024.
- Self-Awareness and Its Relation to Other Parameters of the Clinical Inference Process: Michael J. Chandler*: Department of Psychology, University of Rochester, Rochester, New York 14627.
- Operant Conditioning Principles in the Treatment of a Selectively Mute Child: J. Dennis Nolan* and Connie Pence: Department of Psychology, Ohio State University, 1945 North High Street, Columbus, Ohio 43210.
- Concept Identification and Practice in Brain-Damaged and Process-Reactive Schizophrenic Groups: Oscar A. Parsons* and Helen P. Klein: Department of Psychiatry, University of Oklahoma, 800 Northeast Thirteenth Street, Oklahoma City, Oklahoma 73104.
- Measurement of Appropriate and Unnecessary Help-Seeking Dependent Behavior: Sheldon Cottler,* Robert F. Quilty, and Richard J. Palmer: Department of Psychology, Indiana University, Bloomington, Indiana 47401.
- Is an Abbreviated Form of the WISC Valid for Afro-American Children?: Robert J. Resnick* and Alan D. Entin: Department of Psychiatry, Medical College of Virginia, 1200 East Broad Street, Richmond, Virginia 23219.

* Asterisk indicates for whom address is supplied.

SHOULD WISC SUBJECTS EXPLAIN PICTURE ARRANGEMENT STORIES? ¹

JAMES M. HERRELL

AND

JEFFREY H. GOLLAND ²

Brooke General Hospital, Fort Sam Houston, Texas

New York University School of Medicine

To determine whether requiring WISC Ss to explain their Picture Arrangement (PA) sequences affects PA scores, 78 Ss were given the WISC and asked to tell PA stories, and 57 Ss were given the WISC under standard instructions. Forty-two pairs of Ss were matched for age, Full Scale IQ, and reason for referral. The Ss who told PA stories had an average PA score 2.12 scaled score points higher than Ss who did not tell stories ($p < .01$). Similar differences significant at or beyond the .02 level were found for Ss referred for behavioral-emotional problems, Ss with IQs below 100, and Ss 9 years old or less. Trends ($.05 < p < .10$) were found for Ss referred for academic-intellectual problems and Ss with IQs > 100 . No significant difference was found for Ss > 9 years. Clinicians were urged to be aware of the apparently facilitating effects of Ss explaining PA stories when interpreting the WISC.

In clinical practice, many users of Wechsler tests routinely ask Ss to describe the stories depicted in the Picture Arrangement (PA) series after Ss have arranged the cards in each series. This practice is a deviation from standard procedure, and potentially can alter Ss performance on the PA subtest. In their review of procedural variations in intelligence testing, Sattler and Theye (1967) mentioned no study dealing with this variation, nor were the present authors able to locate such a study in the literature.

The purpose of the present study was to determine what effect asking Ss to describe WISC PA stories has on Ss' PA performance.

METHOD

Subjects

Almost all children referred for psychological evaluation at Brooke General Hospital are administered the age-appropriate Wechsler test. For the present study, the S population consisted of all children completing the WISC from January through July, 1968. There were 135 Ss, 99 males and 36 females, ranging in age from 6 through 15 years.

Referrals tend to be for one of two fairly discrete

reasons. Children are referred either for academic-intellectual problems (A-I) or behavioral-emotional problems (B-E). The appropriate determination is made directly from the referral. A-I Ss are those referred specifically for evaluation of intellectual functioning to meet state requirements for special education or to confirm a diagnosis of retardation or minimal brain dysfunction. In the B-E category are Ss referred for assessment of dynamics and recommendations for treatment. Such Ss represent a variety of psychological disturbances, such as enuresis, uncontrollable acting out, poor peer relations, or withdrawn, depressive behavior. The two authors were in agreement on the appropriate category for all Ss used in the experiment.

Prior to the initiation of this study in April 1968, clinic practice was for A-I referrals to tell PA stories for their projective value; B-E Ss completed more extensive projective tests and were not asked for PA stories. This procedure was altered when the study was begun, resulting in 51 A-I Ss and 27 B-E Ss who told stories and 31 A-I and 26 B-E Ss who did not.

All testing was performed by psychology technicians who were specially trained in the administration and scoring of the tests, and who were unaware of the rationale of the experiment. Stories were obtained by asking S, after each trial, to explain what the PA pictures portrayed. Scoring was independent of these explanations.

RESULTS

Before any difference in PA scores between administrative groups can be attributed to procedural variation, it must be assumed that the two groups would perform similarly if tested under similar conditions, that is, the two groups must have similar IQs. To insure that such was the case, a matching procedure

¹ The authors want to thank the psychology technicians at the hospital, Sam Sherrod, John Stahl, Dennis Gurrie, Bruce Engel, Douglas Abell, James Bowman, and James Wickis, who administered all tests involved, and Roger Gardner who not only administered tests, but helped compile the data.

² At Brooke General Hospital at the time the study was conducted. Requests for reprints should be sent to Jeffrey H. Golland, 145 Fourth Avenue, New York, New York 10003.

TABLE 1
MEAN PA SCORE DIFFERENCES BETWEEN
MATCHED PAIRS

Item	N	M_{Diff}	t	p
All Ss	42	2.12	3.13	<.01
B-E referrals	19	2.00	2.67	<.02
A-I referrals	23	2.18	1.93	<.10
IQ > 100	11	2.09	2.11	<.10
IQ < 100	30	2.20	2.44	<.05
Age < 9	20	1.80	4.13	<.001
Age > 9	21	2.33	1.58	>.10

Note.—Positive M_{Diff} 's indicate higher scores for Ss who explained PA stories.

was used by which Ss were matched based on Full Scale IQ.

The use of post hoc matching was particularly suspect in this case due to the contamination of the matching variable by the dependent variable. The result of such contamination would be to suppress the appearance of a PA difference and bias the experiment in favor of the null hypothesis. For lack of a better control, the experiment was conducted with a conservative bias.

Criteria for matching were three: (a) Age: All Ss were matched with someone within a year of the same age. Age at last birthday was used. (b) Reason for referral: Each was categorized into one of the two types of referrals and matched with someone of that type. (c) Level of intellectual functioning: Each S was matched with someone whose Full Scale IQ was within five points of his own.

Using these criteria, all Ss who could be were matched, 62 males and 22 females, for a total of 42 pairs, and these pairs were used in all comparisons.

The statistics for the unmatched Ss indicated that the method of selection of type of administration was not random. Unmatched Ss who told PA stories had an average IQ nearly seven points higher and PA scores of three points higher than control Ss, emphasizing the necessity for matching.

The statistical test for all comparisons was the t test for correlated samples (McNemar, 1962). Reported probabilities are all for two-tail tests. Table 1 contains mean differences between PA scaled scores of Ss who told stories and those who did not.

The first comparison was for the total

sample of pairs. Mean difference was 2.12 scaled score points, a difference significant at the .01 level. Explaining PA stories seems to facilitate PA performance.

Mean difference for Ss referred for B-E problems was 2.00, significant at the .02 level. Mean differences for A-I Ss was 2.18, with $p < .10$. Even though the magnitude of difference is similar for both groups, the difference seems more reliable for the B-E Ss.

For Ss with Full Scale IQs 100 or greater, the mean difference was 2.09, $p < .10$. Mean difference for Ss with Full Scale IQs below 100 was 2.20, significant at the .05 level, suggesting that the effect is more reliable for the less intelligent Ss.

Mean difference for Ss 9 years or less was 1.80, significant at the .001 level. Mean difference for Ss 10 years or older was 2.33, with $p > .10$. Although the difference is larger for the older children, the effect is more reliable for the younger ones.

DISCUSSION

Although results are equivocal with certain subsamples, it seems clear that having Ss explain PA stories does facilitate performance. Perhaps Ss who learn they must justify their responses are less careless or attend more closely to the task. Whatever the reason, use of this procedural variation with clinical cases tends to increase S's scores by as much as 2 scaled score points, which in turn increases Performance IQ by from 2 to 3 points and Full Scale IQ from 1 to 3 points. Except in those rare cases for which an exact numerical estimate of intelligence is needed, these increments are in themselves insignificant. However, when clinicians engage in profile analysis or scale comparisons, a difference of 2 or 3 points is clinically important. Clinicians who do require PA stories on the WISC should be aware that they are obtaining higher PA scores than if they followed standard procedure, and should interpret findings accordingly.

REFERENCES

- McNEMAR, Q. *Psychological statistics*. New York: Wiley, 1962.
SATTTLER, J. M., & THEYE, F. Procedural, situational, and interpersonal variables in individual intelligence testing. *Psychological Bulletin*, 1967, 68, 347-360.

(Received October 3, 1968)

RELATIONSHIP OF MANIFEST ANXIETY TO REPRESSION-SENSITIZATION ON THE MMPI

PATRICK F. SULLIVAN¹ AND LYNN K. ROBERTS

Veterans Administration Center, Des Moines, Iowa

This study investigated the effect of item commonality on the relationship between the Manifest Anxiety (*MA*) and Repression-Sensitization (*R-S*) scales with respect to an earlier conclusion that the scales have similar psychological meaning. Although it was found that the 29 common items did make a significant contribution, a strong relationship was found when these common items were omitted. It was concluded that *MA* and *R-S* do appear to have similar, if not identical, psychological meaning and that research findings based on *MA* are probably applicable to *R-S*, and vice versa. If so, this finding would lend support to recent studies which indicate that sensitizers perform with more facility than do repressors on certain types of tasks.

A recent study of the relationship between the Extraversion, Manifest Anxiety (*MA*), and Repression-Sensitization (*R-S*) scales reported a correlation of .87 between *MA* and *R-S* (Golin, Herron, Lakota, & Reineck, 1967). Although they acknowledged the possible effects of item overlap, they concluded that these two scales have practically the same psychological meaning. A comparison of the 50-item Taylor *MA* scale (1953) with the 156-item *R-S* scale (Byrne, 1961) indicates a total of 29 items common to both scales. Such extensive overlap might be expected to influence the magnitude of the relationship to a significant degree.

The present study investigated the effects of this item commonality on the relationship between *MA* and *R-S*, since a decreased relationship would question the conclusion of similar psychological meaning. It was predicted that the deletion of the 29 common items would result in a significant decrease in the correlation obtained between *MA* and *R-S*.

METHOD

The *Ss* were 386 males who were evaluated at the Des Moines Veterans Administration Center as applicants for outpatient psychotherapy during the period from 1955 to 1964 and for whom complete MMPIs were available. The data for these *Ss* were their scores on the *R-S* scale, the *MA* scale, the *R-S* scale omitting the 29 common items (*R-S'*), and the *MA* scale minus the 29 common items (*MA'*).

¹ Requests for reprints should be sent to Patrick F. Sullivan, Department of Psychology, Albion College, Albion, Michigan 49224.

The group form numbers of the 29 items omitted from *R-S'* and *MA'* include: 13 (290), 18, 23 (288), 32, 43, 67, 86, 107, 142, 158, 191*, 217, 238, 241*, 242, 263*, 301, 317 (362), 321, 322, 337, 340, 352, 361, 397, 418, 431, 499, 555. (Those items marked with an asterisk (*) are scored when answered True on *MA* and when answered False on *R-S*.) Items 13, 23, and 317 on *MA* are identical to items 290, 288, and 362 on *R-S*.

RESULTS

The mean score of 78.87 ($SD = 23.66$) on *R-S* for this sample is somewhat greater than that reported by Golin et al. (1967). For *MA*, the mean score was 25.96 ($SD = 9.79$). The omission of the 29 common items resulted in a mean decrease of 16.58 ($t = 52.45$, $df = 385$, $p < .001$) for *R-S* and 16.03 ($t = 47.21$, $df = 385$, $p < .001$) for *MA*. Although the average decreases were not significantly different from each other, these figures represent 32% of the *MA* scale as compared to 10.6% of the *R-S* scale.

For this sample, the obtained correlation of .91 ($p < .001$) between *R-S* and *MA* was stronger than that reported by Golin et al. When the common items were excluded from both scales, the relationship diminished significantly ($r = 7.39$, $p < .001$), as expected, but continued to be quite high ($r = .76$, $p < .001$). Similarly, the effect of omitting the common items from only one scale resulted in a diminished but significant degree of relationship between the two scales. The correlations obtained from this latter comparison were $r = .86$ ($p < .001$) for *MA* and *R-S'* and $r = .78$ ($p < .001$) for *R-S* and *MA'*.

DISCUSSION

It seems that although the 29 common items significantly contribute to the relationship between *MA* and *R-S*, there is a high degree of relationship between the scales. These findings would support the conclusion of Golin et al. (1967) that *MA* and *R-S* have similar psychological meaning. Therefore, it seems possible that research findings using the *MA* scale may be applicable to the *R-S* scale and vice versa. If so, the finding that high-*MA* Ss learn better than low-*MA* Ss (Spence, 1960) would tend to support the more recent findings that sensitizers function with more facility than do repressors in stress and ego-involving situations (Hare, 1966; Lazarus & Alfert, 1964; Lomont, 1965; Petzel & Gynther, 1968; Tempone, 1964). Further research may be fruitful to an understanding of the *MA-R-S* relationship.

REFERENCES

- BYRNE, D. The repression-sensitization scale: Rationale, reliability, and validity. *Journal of Personality*, 1961, 29, 334-349.

- GOLIN, S., HERRON, E. W., LAKOTA, R., & REINECK, L. Factor analytic study of the manifest anxiety extraversion, and repression-sensitization scales. *Journal of Consulting Psychology*, 1967, 31, 564-569.
- HARE, R. D. Denial of threat and emotional responses to impending painful stimulation. *Journal of Consulting Psychology*, 1966, 30, 359-361.
- LAZARUS, R. S., & ALFERT, E. Short-circuiting threat by experimentally altering cognitive appraisal. *Journal of Abnormal and Social Psychology*, 1964, 69, 195-205.
- LOMONT, J. F. The repression-sensitization dimension in relation to anxiety responses. *Journal of Consulting Psychology*, 1965, 29, 84-86.
- PETZEL, T. P., & GYNTER, M. D. Task performance of repressors and sensitizers under ego-oriented versus task-oriented instructions. *Journal of Consulting and Clinical Psychology*, 1968, 32, 486-487.
- SPENCE, K. W. *Behavior theory and learning*. Englewood Cliffs, N. J.: Prentice-Hall, 1960.
- TAYLOR, J. A. A personality scale of manifest anxiety. *Journal of Abnormal and Social Psychology*, 1953, 48, 285-290.
- TEMPONE, V. J. Extension of the repression-sensitization hypothesis to success and failure experiences. *Psychological Reports*, 1964, 15, 39-45.

(Received November 1, 1968)